

	12	62	100.0	93	20	AAY41816	E. coli heat-labile
	13	62	100.0	93	20	AAM95526	E. coli heat-labile
	14	62	100.0	93	21	AAY68365	Heat labile toxin
	15	62	100.0	93	22	AAB66239	E. coli verotoxin-1
	16	62	100.0	103	6	AAP50340	Sequence of sub-un
	17	62	100.0	103	17	AAM04857	Synthetic cholera
	18	62	100.0	103	17	AAR94939	Heat labile enteroc
	19	62	100.0	103	17	AAM06606	Cholera toxin B su
	20	62	100.0	103	17	AAM06607	Cholera toxin B su
	21	62	100.0	103	19	AAM80808	Amino acid sequenc
	22	62	100.0	103	22	AAB62365	V. cholera cholerax
	23	62	100.0	103	22	AAB62367	V. cholera cholerax
	24	62	100.0	103	22	AAB62370	V. cholera cholerax
	25	62	100.0	103	22	AAB62379	E. coli LTB proteas
	26	62	100.0	105	22	AAB62369	V. cholera Cholerat
	27	62	100.0	118	11	AAR04163	Cholera Toxin B-su
	28	62	100.0	123	22	AAB62374	E. coli LTB proteas
	29	62	100.0	123	22	AAB62377	E. coli LTB proteas
	30	62	100.0	124	10	AAP93561	B subunit of the h
	31	62	100.0	124	17	AAM06605	Cholera toxin B su
	32	62	100.0	124	19	AAM59770	Amino acid sequenc
	33	62	100.0	124	21	AAY96652	Plant-optimized E.
	34	62	100.0	124	21	AAY96872	Plant-optimized V.
	35	62	100.0	124	22	AAG65992	Cholera toxin B su
	36	62	100.0	124	22	AAB62359	V. cholera strain
	37	62	100.0	124	22	AAB62363	V. cholera cholerat
	38	62	100.0	124	22	AAB62368	V. cholera cholerat
	39	62	100.0	124	22	AAB62373	E. coli LTB proteas
	40	62	100.0	124	22	AAB62375	E. coli LTB proteas
	41	62	100.0	124	22	AAB62376	E. coli LTB proteas
	42	62	100.0	124	22	AAB62378	E. coli LTB proteas
	43	62	100.0	124	22	AAB62380	E. coli LTB proteas
	44	62	100.0	125	22	AAB62372	E. coli LTB proteas
	45	62	100.0	126	12	AAR12630	GtfB_1/CTB chimeri

ALIGNMENTS

RESULT	1	
AAV87461		
ID	AAV87461 standard; peptide; 12 AA.	
AC	AAV87461;	
DT	03-JUL-2000 (first entry)	
DE	Cholera toxin B-enterotoxin B-derived peptide, SEQ ID NO:3.	
KW	Cholera toxin subunit B; CTxB; heat labile enterotoxin subunit B; EtXB; beta-4-alpha-2 loop; GM-1 ganglioside receptor; immunomodulation; adjuvant; immune disorder; diarrhoea.	
OS	Vibrio cholerae.	
OS	Escherichia coli.	
PX	MOZ000141L4-A1.	
PN	16-MAR-2000.	
PD		
PX	07-SEP-1999; 99WO-GH02970.	
PF		
PX	07-SEP-1998; 98GB-0019484.	
PR		
PX	(UYBR-) UNIV BRISTOL.	
PA		
P1	Williams NA, Hirst TR;	
P1		
DR	WPI; 2000-256943/22.	
PX		
PT	Derivatives of Escherichia coli heat labile enterotoxins useful as immunomodulators and for treating diarrhea and which do not bind the glycopolipid receptor GM-1 -	

XX Disclosure; Page 15; 62pp; English.
PS
CC The invention relates to peptide fragments of the Escherichia coli heat
CC labile enterotoxin (Etx) and its closely related homologue, cholera
CC toxin (Ctx) from Vibrio cholerae which do not bind to the ubiquitous
CC GM-1 ganglioside receptors on host cell surfaces. Both Etx and Ctx are
CC composed of one A subunit and five identical B subunits. The A subunit
CC is responsible for toxicity, possessing adenosine diphosphate (ADP) ADP-
CC ribosyltransferase activity, while the B subunits (EtxB and CtxB)
CC facilitate the entry of subunit A into the host cell via the binding and
CC cross-linking of GM-1 receptors. Although GM-1 binding is responsible
CC for some of the effects of Etx and Ctx, it has been found that certain
CC effects of the toxins, such as immunomodulation, are not mediated
CC through GM-1 binding. The peptides of the invention are fragments of the
CC beta-4- alpha-2 loop of EtxB and/or CtxB, exerting the same effects as
CC normal EtxB and CtxB subunits, except that they do not bind or cross link
CC GM-1. They may be used in medicine as an immunomodulator or adjuvant.
CC They may also be used as an inhibitor for toxin-induced diarrhoea.
CC Therefore, the peptides may be used in the production of a composition
CC for treating, preventing and/or modulating a disease associated with an
CC immune disorder and/or toxin-induced diarrhoea. Sequences AAY87460-Y87463
CC represent preferred peptides of the invention, AAY87460 being
CC particularly preferred.
XX
SQ Sequence 12 AA;

Query Match 100.0%; Score 62; DB 21; Length 12;
Best Local Similarity 100.0%; Pred. No. 4.6e-05;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VEVPGSQHIDSQ 12
|||||
DB 1 vevpgsqhidsq 12

RESULT 2
AAP93498
ID AAP93498 standard; protein; 15 AA.
XX
AC AAP93498;

DT 03-MAY-1990 (first entry)

DE CTP3 epitope of the Cholera toxin B subunit.

KW CTP3 epitope of Cholera toxin B subunit; flagellin fusion protein;
KM vaccine; immunotherapy; ds;

PN W08910967-A.

PD 16-NOV-1989.

PF 05-MAY-1989; 89WO-US01932.

PR 05-MAY-1988; 88US-0190570.

PA (PRAX-) PRAXIS BIOLOGICS INC.

PA (STRD) LEYLAND STANDFORD JUNIOR UNIV.

PI Marjarian WR, Stocker BAD, Newton SMC;

DR WPI; 1989-356496/48.

DR N-PSDB; AAN92414.

PT New recombinant flagellin gene including sequence - for heterologous
PT epitope, and expressed fusion proteins, useful in vaccines and for prodn.
PT of antibodies.

PS Disclosure; fig.4B; 137pp; English.

CC This sequence corresponds to the CTP3 epitope of the Cholera toxin B

CC subunit. The DNA sequence encoding this ligates to othersynthetic
CC oligonucleotides to form a new recombinant gene. This encodes
CC a flagellin fusion protein which can be used in vaccines for immuno-
CC therapy.
XX

SQ Sequence 15 AA;

Query Match 100.0%; Score 62; DB 10; Length 15;
Best Local Similarity 100.0%; Pred. No. 5.9e-05;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VEVPGSQHIDSQ 12
|||||
DB 1 vevpgsqhidsq 12

RESULT 3
AAR85125
ID AAR85125 standard; peptide; 15 AA.
XX
AC AAR85125;

DT 13-JUN-1996 (first entry)

DE Cholera toxin B antigenic peptide fragment CTP3.

KW Conjugate; cholera; B toxin; peptide fragment; microparticulate;

KW inert carrier; modified silica; thyroglobulin; oral vaccine;

KW immunisation; infection; insoluble; digestive tract; antigen;

KW intestines; antibodies; secretory; IgA class.

OS Vibrio cholerae.

PN W09529701-A1.

PD 09-NOV-1995.

PF 02-MAY-1995; 95WO-EP01661.

PR 03-MAY-1994; 94TL-0109519.

PA (YEDA) YEDA RES & DEV CO LTD.

PI Marks RS, Mirelman D, Sela M;

DR WPI; 1995-403805/51.

PT Vaccines for oral immunisation against infecting agents, e.g.

PT cholera - comprise a conjugate of an antigen of an infecting agent

PT covalently bound to micro:particulate inert carrier, e.g. modified

PT aldehyde silica

PS Claim 7; Page 25; 40pp; English.

XX A compsn. comprising a conjugate of an antigenic cholera B toxin

CC peptide fragment (e.g. AAR85126-30 and esp. AAR85125), covalently

CC bound to a microparticulate inert carrier (e.g. modified silica or

CC thyroglobulin) can be used as an oral vaccine for immunisation

CC against cholera infection. The inert carrier is insoluble in the

CC digestive tract, allowing presentation of the antigen in the

CC intestines, where it will elicit antibodies mainly of the

CC secretory IgA class.

SQ Sequence 15 AA;

Query Match 100.0%; Score 62; DB 16; Length 15;
Best Local Similarity 100.0%; Pred. No. 5.9e-05;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VEVPGSQHIDSQ 12
|||||

ID AAR76748 standard; Protein; 23 AA.
 XX AAR76748;
 AC
 XX
 XX 18-MAR-1996 (first entry)
 DT
 XX Residues 50-64 of cholera toxin B subunit and FimH 224-226.
 DE
 XX FimH; type 1 fimbriae; organelle; adhesin; alpha-D-mannoside residue;
 KW FimH; FimH; receptor binding site; PCR; amplify; ss.
 XX
 OS Chimeric - Vibrio cholerae.
 XX Chimeric - Escherichia coli.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 1
 FT Peptide /note= "Represents FimH residue 224"
 FT Peptide 2..4
 FT Peptide /note= "Linker peptide"
 FT Peptide 5..19
 FT Peptide /note= "Cholera toxin B subunit 50-64"
 FT Peptide 20..22
 FT Misc-difference 23 /note= "Linker peptide"
 FT Misc-difference /note= "Represents FimH residue 226"
 XX
 XX W09520657-A1.
 PN
 XX 03-AUG-1995.
 PD
 XX 27-JAN-1995; 95WO-DK00042.
 PF
 XX 27-JAN-1994; 94US-0187166.
 PR
 XX (GXRI-) GX BIOSYSTEMS AS.
 PA
 XX Hasty DL, Klemm P, Molin S, Pallesen L, Sokurenko EV;
 XX
 XX WPI; 1995-275442/36.
 DR N-PSDB; AAQ93061.
 XX
 XX Receptor specific bacterial adhesins - useful for targeting active
 PT compounds and microbial cells to locations of receptors
 PS
 XX Example 3; Page 58; 152pp; English.
 CC This sequence is encoded by a fragment of the the plasmid pUPA93
 CC which was used in the production of FimH fusion genes comprising
 CC the cholera toxin B subunit inserted into the fimH gene. This insert
 CC shows the inclusion of the B subunit into the FimH protein at position
 CC 224-226. The chimeric genes were then opt. further modified by insertion
 CC of the hepatitis B virus surface antigen pre-S2 region into a different
 CC position of the FimH adhesin of type 1 fimbriae. Restriction site handles
 CC (BglII-sites) were introduced into the fimH gene, and the foreign
 CC epitopes are then inserted in-frame. In the selected positions the
 CC insertion of the epitopes did not significantly alter the adhesive
 CC function of the FimH protein. The expression of the chimeric proteins
 CC on the surface of fimbriae on bacterial hosts illustrated the possibility
 CC of using bacterial adhesins as general presenters of foreign antigens and
 CC epitopes. These chimeric genes may be used in the production of variant
 CC FimH adhesins which may be useful for targeting active compounds
 CC and microbial cells to locations comprising selected receptors to which
 CC the adhesins bind.
 XX
 SQ Sequence 23 AA;

Query Match 100.0%; Score 62; DB 16; Length 23;
 Best Local Similarity 100.0%; Pred. No. 9.6e-05;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 VEVPSQHDSQ 12
 |||
 f

DB 5 vevpsqhsdq 16
 RESULT 7
 ID AAP30265 standard; Protein; 26 AA.
 XX AAP30265;
 AC
 XX 21-APR-1992 (first entry)
 DT
 XX Sequence of amino acids 50-75 of the cholera toxin B1 subunit which
 DE carries an Arg at posns. 67 and 73.
 XX
 KW Cholera vaccine; therapy; E.coli infection; enterotoxin LT.
 XX
 OS Vibrio cholerae.
 XX
 XX EP95426-A.
 PN
 XX 30-NOV-1983.
 PD
 XX 26-MAY-1983; 83BP-0401052.
 PF
 XX 26-MAY-1982; 82FR-0009167.
 PR
 XX (CNRS) CNRS CENT NAT RECH SCI.
 PA (INSP) INST PASTEUR.
 XX
 PI Milhaud G, Raulais D, Rivaille P, Siffert O, Dodin A;
 PI Guyongruaz A, Delmas A;
 XX
 XX WPI; 1983-834645/49.
 DR
 XX Cholera toxin B, sub-unit polypeptide(s) as vaccines and
 PT medicaments - effective against Escherichia coli and Vibrio
 PT cholerae infections, are prepd. by solid phase peptide synthesis
 XX
 PS Claim 7; Page 11; 13pp; French.
 CC The inventors claim cholera toxin B1 subunit sequences which carry
 CC Arg in posn. 35, 67 or 73 (AAP30264, AAP30265, AAP30600). Tests have
 CC shown that Arg in posns. 35, 67 and 73 play an important role in fixing
 CC these toxins to cell walls. The peptides are used in the treatment of,
 CC and vaccination against cholera infections and animal and human
 CC infections due to E. coli (enterotoxin LT). The medicament may be
 CC administered by oral, intraperitoneal, sub-cutaneous or intravenous
 CC routes. For vaccines, pref. peptides having 15-30 AAs are injected
 CC without carriers. Unit dose when used as a medicament is 50-500mg as
 CC a vaccine 1-10mg of active cpd.
 XX
 SQ Sequence 26 AA;

Query Match 100.0%; Score 62; DB 4; Length 26;
 Best Local Similarity 100.0%; Pred. No. 0.00011;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 VEVPSQHDSQ 12
 |||
 Db 1 vevpsqhsdq 12
 RESULT 8
 ID AAP50439 standard; protein; 41 AA.
 AAP50439
 XX
 AC AAP50439;
 XX
 DT 01-JAN-1980 (first entry)
 XX
 DE Network polymer which comprises a series of composite E. coli heat-
 labile toxin (LT)/heat-stable toxin (ST) polypeptide repeating units.

XX Heat-labile enterotoxin; heat-stable toxin; vaccine; ss.
 KM Synthetic.
 OS
 XX WO8502611-A.
 PN
 XX 20-JUN-1985.
 PD
 XX 12-DEC-1984; 84WO-US02030.
 PF
 XX 12-DEC-1983; 83US-0559469.
 PR
 XX (SCRI-) SCRIPPS CLINIC RES.
 PA
 XX Houghten RA;
 PI
 XX WPI; 1985-159230/26.
 DR
 XX
 XX New synthetic polypeptide(s) for part of E. coli enterotoxin(s) -
 PT useful for vaccination of mammals against the enterotoxin(s)
 PS
 XX Claim 8; Page 100; 120pp; English.
 PS
 XX The repeating units are bonded together by intramolecular
 CC interpolypeptide cystine bonds formed between oxidized Cys residues
 CC of the repeating units. This polypeptide may be used in the
 CC vaccination of mammals for protection against the enterotoxins. The
 CC composite polypeptide is made by solid phase synthesis or
 CC recombinant DNA techniques. X9 is Met or Lys, X12 is Ala or Glu,
 CC X39 is Asn or Tyr and X46 is Tyr or Asn. See also AAP50439-57.
 CC
 XX
 SQ Sequence 41 AA;

Query Match 100.0%; Score 62; DB 6; Length 41;
 Best Local Similarity 100.0%; Pred. No. 0.00018;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VEVPGSQHDSQ 12
 |||||||||
 Db 14 vevpgsqhdsq 25

RESULT 9
 AAP50436
 ID AAP50436 standard; protein; 46 AA.
 XX
 AC AAP50436;
 XX
 DT 01-JAN-1980 (first entry)
 DT
 XX Network polymer which comprises a series of composite E. coli heat-
 DE labile toxin (LT)/heat-stable toxin (ST) polypeptide repeating units.
 DE
 XX Heat-labile enterotoxin; heat-stable toxin; vaccine; ss.
 KM
 XX Synthetic.
 OS
 XX WO8502611-A.
 PN
 XX 20-JUN-1985.
 PD
 XX 12-DEC-1984; 84WO-US02030.
 PF
 XX 12-DEC-1983; 83US-0559469.
 PR
 XX (SCRI-) SCRIPPS CLINIC RES.
 PA
 XX Houghten RA;
 PI
 XX WPI; 1985-159230/26.
 DR
 XX

PT New synthetic polypeptide(s) for part of E. coli enterotoxin(s) -
 PT useful for vaccination of mammals against the enterotoxin(s)
 PS
 XX Claim 8; Page 100; 120pp; English.
 PS
 XX The repeating units are bonded together by intramolecular
 CC interpolypeptide cystine bonds formed between oxidized Cys residues
 CC of the repeating units. This polypeptide may be used in the
 CC vaccination of mammals for protection against the enterotoxins. The
 CC composite polypeptide is made by solid phase synthesis or
 CC recombinant DNA techniques. X9 is Met or Lys, X12 is Ala or Glu,
 CC X39 is Asn or Tyr and X46 is Tyr or Asn. See also AAP50439-57.
 CC
 XX
 SQ Sequence 46 AA;

Query Match 100.0%; Score 62; DB 6; Length 46;
 Best Local Similarity 100.0%; Pred. No. 0.00021;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VEVPGSQHDSQ 12
 |||||||||
 Db 16 vevpgsqhdsq 27

RESULT 10
 AAP30600
 ID AAP30600 standard; protein; 47 AA.
 XX
 AC AAP30600;
 XX
 DT 21-APR-1992 (first entry)
 DT
 XX Sequence of amino acids 350-75 of the cholera toxin B1 subunit which
 DE carries an Arg at posns. 35, 67 and 73.
 DE
 XX Cholera vaccine; therapy; E.coli infection; enterotoxin LT.
 KM
 XX Vibrio cholerae.
 OS
 XX EP95426-A.
 PN
 XX 30-NOV-1983.
 PD
 XX 26-MAY-1983; 83EP-0401052.
 PF
 XX 26-MAY-1982; 82FR-0009167.
 PR
 XX (CNRS) CNRS CENT NAT RECH SCI.
 PA (INSP) INST PASTEUR.
 XX
 PI Milhaud G, Raulais D, Rivaille P, Siffert O, Dodin A;
 PI Guyongruaz A, Delmas A;
 PI WPI; 1983-834645/49.
 DR
 XX Cholera toxin B, sub-unit polypeptide(s) as vaccines and
 PT medicaments - effective against Escherichia coli and Vibrio
 PT cholerae infections, are prepd. by solid phase peptide synthesis
 PT
 XX Claim 8; Page 11; 13pp; French.
 PS
 XX The inventors claim cholera toxin B1 subunit sequences which carry
 CC Arg in posn. 35, 67 or 73 (AAP30264, AAP30265, AAP30600). Tests have
 CC shown that Arg in posns. 35, 67 and 73 play an important role in fixing
 CC these toxins to cell walls. The peptides are used in the treatment of,
 CC and vaccination against, cholera infections and animal and human
 CC infections due to E. coli (enterotoxin LT). The medicament may be
 CC administered by oral, intraperitoneal, sub-cutaneous or intravenous
 CC routes. For vaccines, pref. peptides having 15-30 AAs are injected
 CC without carriers. Unit dose when used as a medicament is 50-500mg as
 CC a vaccine 1-10mg of active cpd.
 CC

SQ Sequence 47 AA;

Query Match 100.0%; Score 62; DB 4; Length 47;

Best Local Similarity 100.0%; Pred. No. 0.00021;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;QY 1 VEVPGSQHDSQ 12
| | | | | | | | | |
DB 22 vevpgsqhdsq 33

RESULT 11

AAR72545
ID AAR72545 standard; peptide; 93 AA.

XX AAR72545;

XX 28-NOV-1995 (first entry)

XX ADP-ribosylating toxin (verotoxin-1 B-subunit).

XX ADP-ribosylating toxin; pertussis holotoxin; B-subunit;

XX active site; E. coli heat labile toxin; verotoxin-1;

XX Bordetella pertussis vaccines.

XX Bacteria sp.

XX EP646599-A.

XX 05-APR-1995.

XX 23-AUG-1994; 94EP-0306219.

XX 24-AUG-1993; 93US-0110947.

XX 31-MAY-1994; 94US-0251121.

XX (CONN-) CONNAUGHT LAB LTD.

XX (UYAL-) UNIV ALBERTA.

XX Armstrong GD, Cockle SA, Hazes B, Klein MH, Loosmore S;

XX Oomen R, Read RJ, Stein PE;

XX WPI; 1995-132623/18.

XX New modified forms of pertussis holotoxin - developed using

XX crystalline forms of pertussis holotoxin and its complexes with

XX other molecules

XX Disclosure; Fig 5; 54pp; English.

XX AAR72540-R72545 are structurally equivalent B-subunits from three

XX ADP-ribosylating toxins, pertussis holotoxin (PT), E. coli heat

XX labile toxin (LT), and verotoxin-1 (VT). The structural

XX information obtained from these comparisons was used to identify

XX sites which contribute to PT's biological activity. By modifying

XX these sites the claimed PT mutants of the invention were produced,

XX they can be used in the development of vaccines against Bordetella

XX pertussis infection.

XX SQ Sequence 93 AA;

AA41816
ID AA41816 standard; peptide; 93 AA.

XX AA41816;

XX 08-DEC-1999 (first entry)

XX Escherichia coli verotoxin-1 B-subunit.

XX ADP-ribosylating toxin; PT: crystalline pertussis holotoxin; ETA; DT;

XX three-dimensional structure; LT; immunoprotective; infection.

XX Escherichia coli.

XX US5965385-A.

XX 12-OCT-1999.

XX 06-JUN-1995; 95US-0467974.

XX 22-AUG-1994; 94US-0292968.

XX 24-AUG-1993; 93US-0110947.

XX 31-MAY-1994; 94US-0251121.

XX (CONN-) CONNAUGHT LAB LTD.

XX (UYAL-) UNIV ALBERTA.

XX Read RJ, Stein PE, Cockle SA, Loosmore S, Klein MH, Armstrong GD;

XX Hazes B, Oomen RP;

XX WPI; 1999-579908/49.

XX New method for producing modified pertussis holotoxin -

XX Example 3; Fig 5; 41pp; English.

XX A method has been developed of producing a modified pertussis holotoxin,

XX involving analysis of the 3-dimensional form of the crystalline

XX holotoxin. The pertussis holotoxin modification process comprises:

XX (1) identification of at least one amino acid (aa) residue of the

XX holotoxin for modification by analysing the 3-dimensional form of the

XX crystalline holotoxin, in relation to known information of the protein

XX structure and function; (2) effecting mutagenesis (by removing or

XX replacing a nucleotide sequence encoding at least one (aa) of a tox

XX operon; and (3) expressing mutant tox box in a Bordetella organism to

XX produce the modified holotoxin. This method is used for modifying

XX pertussis holotoxin, by studying its 3-dimensional crystalline

XX structure. Modifying the holotoxin, alters its biological properties.

XX By analysing the 3-dimensional crystalline structure of the pertussis

XX holotoxin, functional (aa) which affect biological properties of the

XX pertussis holotoxin can be identified. This can be used to predict (aa)

XX which contribute to the toxicity of the holotoxin to produce

XX immunoprotective, genetically-detoxified analogues of pertussis

XX holotoxin. The present sequence represents an ADP-ribosylating toxin

XX invention.

SQ Sequence 93 AA;

Query Match 100.0%; Score 62; DB 20; Length 93;

Best Local Similarity 100.0%; Pred. No. 0.00045;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;QY 1 VEVPGSQHDSQ 12
| | | | | | | | | |
DB 40 vevpgsqhdsq 51

RESULT 13

AAW95226
ID AAW95226 standard; peptide; 93 AA.

AC AAW95226;
 XX
 DT 16-MAR-1999 (first entry)
 XX
 DE E. coli heat-labile toxin (LT) beta-subunit sequence.
 XX
 KM Pertussis holotoxin; PT; modified; effector; toxicity; cell binding;
 KM enzyme; ADP-ribosylation; toxin; crystalline; X-ray crystallography;
 KM structural analysis; interacting site; mitogenicity; adjuvanticity;
 KM heat-labile; LT.
 XX
 OS Escherichia coli.
 XX
 PN US5856122-A.
 XX
 PD 05-JAN-1999.
 XX
 PF 22-AUG-1994; 94US-0292968.
 XX
 PR 22-AUG-1994; 94US-0292968.
 PR 24-AUG-1993; 93US-0110947.
 PR 31-MAY-1994; 94US-0251121.
 XX
 PA (UYAL-) UNIV ALBERTA.
 XX
 PI Armstrong GD, Cockle SA, Hazes B, Klein MH, Loosmore S;
 PI Oomen RP, Read RJ, Stein PE;
 DR WPI; 1999-105104/09.
 XX
 PT Modifications to e.g. enzymatic activity, mitogenicity and cell
 PT binding of pertussis holotoxin - by identifying interaction sites of
 PT a molecule with crystalline toxin and modifying the identified site
 XX
 PS Example 3; Fig 5; 40pp; English.
 XX
 CC The invention relates to methods of preparing a pertussis holotoxin (PT)
 CC having a modified biological activity. One method comprises identifying
 CC at least 1 site in a PT that interacts with a molecule that is capable of
 CC forming a complex with the holotoxin and which molecule is an effector
 CC molecule which is an adenine nucleotide and which site contributes to
 CC toxicity, cell binding or enzymatic activity of PT. The functional
 CC interacting site(s) are identified by analysing the three dimensional
 CC structure of crystalline PT, determined by X-ray crystallography. The
 CC identified interacting site(s) are modified to alter toxicity, cell
 CC binding or enzyme activity of the PT. The methods can be used to alter a
 CC biological activity such as toxicity, enzymatic activity, mitogenicity,
 CC cell binding and adjuvanticity of the PT. The three-dimensional structure
 CC of PT have functional and/or structural resemblance to other bacterial
 CC toxins such as diphtheria toxin (DT), Pseudomonas exotoxin A (ETA), the
 CC heat-labile toxin of E. coli (LT) and verotoxin-1 (VT). The present
 CC sequence represents the beta-subunit of LT toxin.
 XX
 SQ Sequence 93 AA;

Query Match 100.0%; Score 62; DB 20; Length 93;
 Best Local Similarity 100.0%; Pred. No. 0.00045;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VEVPSQHDSQ 12
 DB 40 vevpsqhdsq 51
 |||||||||

RESULT 14
 ID AAY68365 standard; Peptide; 93 AA.
 XX
 AC AAY68365;
 XX
 DT 17-APR-2000 (first entry)
 XX

DE Heat labile toxin B subunit SEQ ID NO:26.
 XX
 KM Pertussis toxin; heat labile toxin; verotoxin; exotoxin A; holotoxin;
 KM diphtheria toxin; ADP-ribosylating toxin; mannose binding protein;
 KM infection; crystal structure; X-ray crystallography; detoxification;
 KM immunogenic.
 XX
 OS Escherichia coli.
 XX
 PN US6018022-A.
 XX
 PD 25-JAN-2000.
 XX
 PF 06-JUN-1995; 95US-0467976.
 XX
 PR 22-AUG-1994; 94US-0292968.
 PR 24-AUG-1993; 93US-0110947.
 PR 31-MAY-1994; 94US-0251121.
 XX
 PA (CONN-) CONNAUGHT LAB LTD.
 PA (UYAL-) UNIV ALBERTA.
 XX
 PI Read RJ, Stein PE, Cockle SA, Loosmore S, Klein MH, Armstrong GD;
 PI Hazes B, Oomen RP;
 DR WPI; 2000-136703/12.
 XX
 PT Crystalline form of isolated pertussis holotoxin useful in studying
 PT proteins which have functional resemblance -
 XX
 PS Example 3; Fig 5; 42pp; English.
 XX
 CC The present invention describes a crystalline form of isolated
 CC pertussis holotoxin, in which the molecules of pertussis toxin have
 CC a three dimensional structure represented in the specification,
 CC complexed with a polysaccharide molecule capable of forming a complex
 CC with the holotoxin. The crystalline form of the pertussis holotoxin
 CC can be used in a comparison with other proteins which have functional
 CC resemblance to pertussis holotoxin with the aim of modifying other
 CC proteins. Identifying the unknown sites of toxicity by comparison
 CC with the three dimensional structure of pertussis holotoxin provides a
 CC technique for detoxification of toxins to produce useful immunogenic
 CC but non-toxic analogues. It can also be used as a primary standard to
 CC measure the quantity, purity or efficacy of less pure compositions
 CC containing pertussis toxin. AAY68340 to AAY68385 represent peptides
 CC used in the exemplification of the present invention.
 XX
 SQ Sequence 93 AA;

Query Match 100.0%; Score 62; DB 21; Length 93;
 Best Local Similarity 100.0%; Pred. No. 0.00045;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VEVPSQHDSQ 12
 DB 40 vevpsqhdsq 51
 |||||||||

RESULT 15
 ID AAB66239 standard; Protein; 93 AA.
 XX
 AC AAB66239;
 XX
 DT 03-APR-2001 (first entry)
 XX
 DE E coli verotoxin-1 B subunit SEQ ID NO: 26.
 XX
 KM Pertussis toxin; crystal structure; whooping cough; biological activity;
 KM lymphocytosis-promoting factor; histamine-sensitising factor;
 KM islet-activating protein.
 XX

OS Escherichia coli.

XX US6168928-B1.

XX 02-JAN-2001.

XX 21-MAY-1998; 98US-0082514.

XX 22-AUG-1994; 94US-0292968.

PR 24-AUG-1993; 93US-0110947.

PR 31-MAY-1994; 94US-0251121.

XX (CONN-) CONNAUGHT LAB LTD.

XX Read RJ, Cockle SA, Oomen RP, Loosmore S, Klein MH, Armstrong GD;

PI Hazes B, Stein PE;

XX WPI; 2001-122260/13.

DR Modifying pertussis holotoxin to produce detoxified PT analogs,

XX PT comprising analyzing crystalline structure of toxin, to identify sites

PT of toxicity, cell binding or enzyme activity of PT and modifying

PT identified site -

XX Example 3; Fig 5; 41pp; English.

PS The present invention provides a method for producing a pertussis toxin

XX (also designated lymphocytosis-promoting factor, histamine-sensitising

CC factor and islet activating protein) with a modified biological activity,

CC involving analysing the crystal structure of the protein to identify

CC active sites which can then be modified. This may lead to an alteration

CC in the toxicity, cell binding or enzyme activity of the toxin. This can

CC be used in the production of immunoprotective analogues of pertussis

CC toxin. Pertussis toxin is the cause of whooping cough following infection

CC by Bordetella pertussis.

XX SQ Sequence 93 AA;

Query Match 100.0%; Score 62; DB 22; Length 93;

Best Local Similarity 100.0%; Pred. No. 0.00045;

Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VEYPSQSQTDSQ 12

IIIIIIIIIIII

DB 40 VEYPSQSQTDSQ 51

Search completed: July 3, 2002, 08:54:57

Job time: 287 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 3, 2002, 08:55:38 ; Search time 34.88 Seconds
(without alignments)
8.403 Million cell updates/sec

Title: US-09-786-648-3
Perfect score: 62
Sequence: 1 VEVPGSQHDSQ 12

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 08
Maximum Match 1008
Listing first 45 summaries

Database : Issued Patents AA:*

- 1: /cgcn2_6/ptodata/2/1aa/5A.COMB.pep:*
- 2: /cgcn2_6/ptodata/2/1aa/5B.COMB.pep:*
- 3: /cgcn2_6/ptodata/2/1aa/6A.COMB.pep:*
- 4: /cgcn2_6/ptodata/2/1aa/6B.COMB.pep:*
- 5: /cgcn2_6/ptodata/2/1aa/6CTUS.COMB.pep:*
- 6: /cgcn2_6/ptodata/2/1aa/6ackfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	62	100.0	93	2	US-08-292-968-26 Sequence 26, Appl
2	62	100.0	93	2	US-08-467-974-26 Sequence 26, Appl
3	62	100.0	93	2	US-08-467-536-26 Sequence 26, Appl
4	62	100.0	93	3	US-08-467-976-26 Sequence 26, Appl
5	62	100.0	93	4	US-09-082-514-26 Sequence 26, Appl
6	62	100.0	102	3	US-08-952-337-5 Sequence 5, Appl
7	62	100.0	102	3	US-08-952-337-6 Sequence 6, Appl
8	62	100.0	103	2	US-08-472-171-2 Sequence 2, Appl
9	62	100.0	103	2	US-08-894-526-2 Sequence 2, Appl
10	62	100.0	103	2	US-09-013-047-2 Sequence 2, Appl
11	62	100.0	103	4	US-09-374-597-2 Sequence 2, Appl
12	62	100.0	103	4	US-09-191-852-21 Sequence 21, Appl
13	62	100.0	103	5	PCT-US95-13376-21 Sequence 21, Appl
14	62	100.0	123	3	US-08-952-337-1 Sequence 1, Appl
15	62	100.0	123	3	US-08-952-337-2 Sequence 2, Appl
16	62	100.0	124	2	US-08-747-410-2 Sequence 2, Appl
17	62	100.0	371	1	US-08-829-026A-6 Sequence 6, Appl
18	56	90.3	124	1	US-08-449-045C-4 Sequence 4, Appl
19	56	90.3	124	2	US-08-435-605A-12 Sequence 12, Appl
20	56	90.3	124	6	5223610-3 Patent No. 5223610
21	39	62.9	346	2	US-08-602-359A-34 Sequence 34, Appl
22	37	59.7	448	4	US-08-878-989-2 Sequence 2, Appl
23	37	59.7	448	4	US-09-272-796-2 Sequence 2, Appl
24	37	59.7	458	4	US-09-344-700-4 Sequence 4, Appl
25	36	58.1	855	4	US-09-027-337-2 Sequence 2, Appl
26	34	54.8	93	4	US-09-069-023-15 Sequence 15, Appl
27	34	54.8	219	4	US-09-069-023-15 Sequence 12, Appl

28	34	54.8	459	6	5194375-6 Patent No. 5194375
29	34	54.8	775	2	US-08-714-070A-1 Sequence 1, Appl
30	33	53.2	321	3	US-09-039-609-4 Sequence 4, Appl
31	33	53.2	458	3	US-09-039-609-2 Sequence 2, Appl
32	33	53.2	646	4	US-09-625-188-10 Sequence 10, Appl
33	33	53.2	845	1	US-08-416-950-11 Sequence 11, Appl
34	33	53.2	845	2	US-08-469-830-11 Sequence 11, Appl
35	32	51.6	282	1	US-07-672-304-3 Sequence 3, Appl
36	32	51.6	297	1	US-08-011-398B-4 Sequence 4, Appl
37	32	51.6	297	1	US-08-464-051-4 Sequence 4, Appl
38	32	51.6	297	2	US-08-462-498-4 Sequence 4, Appl
39	32	51.6	297	3	US-08-554-385-3 Sequence 3, Appl
40	32	51.6	321	4	US-09-171-461-22 Sequence 22, Appl
41	32	51.6	360	1	US-08-674-612-3 Sequence 3, Appl
42	32	51.6	360	1	US-08-469-421-14 Sequence 14, Appl
43	32	51.6	360	1	US-08-250-975-14 Sequence 14, Appl
44	32	51.6	360	2	US-08-920-296-3 Sequence 3, Appl
45	32	51.6	360	2	US-08-605-002A-14 Sequence 14, Appl

ALIGNMENTS

RESULT 1
US-08-292-968-26
Sequence 26, Application US/08292968
Patent No. 5856122

GENERAL INFORMATION:

APPLICANT: READ, Randy J.
APPLICANT: COCKLE, Stephen E.
APPLICANT: OOMEN, Raymond P.
APPLICANT: LOOSMORE, Sheena
APPLICANT: KLEIN, Michel H.
APPLICANT: ARMSTRONG, Glen D.

TITLE OF INVENTION: MODIFICATION OF PERTUSSIS TOXIN
NUMBER OF SEQUENCES: 46
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sim & McBurney
STREET: Suite 701, 330 University Avenue
City: Toronto
STATE: Ontario
COUNTRY: Canada
ZIP: M5G 1R7

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/292,968
FILING DATE:

CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/110,947
FILING DATE: 24-AUG-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/251,121
FILING DATE: 31-MAY-1994

ATTORNEY/AGENT INFORMATION:
NAME: STEWART, Michael I.
REGISTRATION NUMBER: 24,973
REFERENCE/DOCKET NUMBER: 1038-388
TELECOMMUNICATION INFORMATION:
TELEPHONE: (416) 595-1155
TELEFAX: (416) 595-1163

INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 93 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear

US-08-292-968-26

Query Match 100.0%; Score 62; DB 2; Length 93;
Best Local Similarity 100.0%; Pred. No. 0.00032;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VEVPGSQHDSQ 12
|||||
DB 40 VEVPGSQHDSQ 51

RESULT 2

US-08-467-974-26
; Sequence 26, Application US/08467974
; Patent No. 5965385
; GENERAL INFORMATION:
; APPLICANT: READ, Randy J.
; APPLICANT: STEIN, Penelope E.
; APPLICANT: COCKLE, Stephen A.
; APPLICANT: OOMEN, Raymond P.
; APPLICANT: LOOSMORE, Sheena
; APPLICANT: KLEIN, Michel H.
; APPLICANT: ARMSTRONG, Glen D.
; APPLICANT: HAZES, Bart
; TITLE OF INVENTION: MODIFICATION OF PERTUSSIS TOXIN
; NUMBER OF SEQUENCES: 46
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: Suite 701, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/467,974
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/467,536
; FILING DATE: 06-JUN-1995
; APPLICATION NUMBER: US 08/292,968
; FILING DATE: 22-AUG-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/251,121
; FILING DATE: 31-MAY-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/110,947
; FILING DATE: 24-AUG-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: STEWART, Michael I.
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-454 MIS:vg
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; INFORMATION FOR SEQ ID NO: 26:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 93 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-467-974-26

Query Match 100.0%; Score 62; DB 2; Length 93;
Best Local Similarity 100.0%; Pred. No. 0.00032;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VEVPGSQHDSQ 12
|||||
DB 40 VEVPGSQHDSQ 51

RESULT 3

US-08-467-536-26
; Sequence 26, Application US/08467536
; Patent No. 597304
; GENERAL INFORMATION:
; APPLICANT: READ, Randy J.
; APPLICANT: STEIN, Penelope E.
; APPLICANT: COCKLE, Stephen A.
; APPLICANT: OOMEN, Raymond P.
; APPLICANT: LOOSMORE, Sheena
; APPLICANT: KLEIN, Michel H.
; APPLICANT: ARMSTRONG, Glen D.
; APPLICANT: HAZES, Bart
; TITLE OF INVENTION: MODIFICATION OF PERTUSSIS TOXIN
; NUMBER OF SEQUENCES: 46
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: Suite 701, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/467,536
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/292,968
; FILING DATE: 22-AUG-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/251,121
; FILING DATE: 31-MAY-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/110,947
; FILING DATE: 24-AUG-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: STEWART, Michael I.
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-455 MIS:vg
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; INFORMATION FOR SEQ ID NO: 26:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 93 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-467-536-26

Query Match 100.0%; Score 62; DB 2; Length 93;
Best Local Similarity 100.0%; Pred. No. 0.00032;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VEVPGSQHDSQ 12
|||||
DB 40 VEVPGSQHDSQ 51

RESULT 4

US-08-467-976-26

```

: Sequence 26 Application US/08467976
: Patent No. 6018022
:
: GENERAL INFORMATION:
: APPLICANT: READ, Randy J.
: APPLICANT: STEIN, Penelope E.
: APPLICANT: COCKLE, Stephen A.
: APPLICANT: COMEN, Raymond P.
: APPLICANT: LOOSMORE, Sheena
: APPLICANT: KLEIN, Michel H.
: APPLICANT: ARMSTRONG, Glen D.
:
: TITLE OF INVENTION: MODIFICATION OF PERTUSSIS TOXIN
: NUMBER OF SEQUENCES: 46
:
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Slim & Mcburney
: STREET: Suite 701, 330 University Avenue
: CITY: Toronto
: STATE: Ontario
: COUNTRY: Canada
:
: ZIP: M5G 1R7
:
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent In Release #1.0, Version #1.25
:
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/467,976
: FILING DATE: 06-JUN-1995
: CLASSIFICATION: 435
:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/292,968
: FILING DATE: 22-AUG-1994
:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/251,121
: FILING DATE: 31-MAY-1994
:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/110,947
: FILING DATE: 24-AUG-1993
:
: ATTORNEY/AGENT INFORMATION:
: NAME: STEWART, Michael I.
: REGISTRATION NUMBER: 24,973
: REFERENCE/DOCKET NUMBER: 1038-453 MIS:vg
: TELEPHONE: (416) 595-1155
: TELEFAX: (416) 595-1163
:
: INFORMATION FOR SEQ ID NO: 26:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 93 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
:
: US-08-467-976-26
:
:
: Query Match 100.0%; Score 62; DB 3; Length 93;
: Best Local Similarity 100.0%; Pred. No. 0.00032;
: Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
:
: Oy 1 VEVPSQHDSQ 12
: Db 40 VEVPSQHDSQ 51
:
: RESULT 5
: US-09-082-514-26
: Sequence 26, Application US/09082514
: Patent No. 6168928
:
: GENERAL INFORMATION:
: APPLICANT: READ, Randy J.
: APPLICANT: STEIN, Penelope E.
: APPLICANT: COCKLE, Stephen A.
: APPLICANT: COMEN, Raymond P.
: APPLICANT: KLEIN, Michel H.

```

APPLICANT: ARMSTRONG, Glen D.
APPLICANT: HAZES, Bat
TITLE OF INVENTION: MODIFICATION OF PERTUSSIS TOXIN
NUMBER OF SEQUENCES: 46
CORRESPONDENCE ADDRESS:
ADDRESSEE: Slim & McBurney
STREET: 6th Floor, 330 University Avenue
CITY: Toronto
STATE: Ontario
COUNTRY: Canada
ZIP: M5G 1R7
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/082,514
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/292,968
FILING DATE: 24-AUG-1994
ATTORNEY/AGENT INFORMATION:
NAME: STEWART, Michael I.
REGISTRATION NUMBER: 24,973
REFERENCE/DOCKET NUMBER: 1038-810
TELECOMMUNICATION INFORMATION:
TELEPHONE: (416) 595-1155
TELEFAX: (416) 595-1163
INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 93 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: Linear
US-09-082-514-26

```

Query Match          100.0%; Score 62; DB 4; Length 93;
Best Local Similarity 100.0%; Pred. No. 0.00037;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      1 VEPESQHDSQ 12
        |||
Db       40 VEPESQHDSQ 51

RESULT   6
US-08-952-337-5
; Sequence 5, Application US/08952337
; Patent No. 6019973
; GENERAL INFORMATION:
; APPLICANT: Holmgren, Jan
; APPLICANT: Lebens, Michael R.
; TITLE OF INVENTION: HYBRID MOLECULES BETWEEN HEAT-LABILE
; FILE REFERENCE: 3846/OD758
; CURRENT APPLICATION NUMBER: US/08/952.337
; CURRENT FILING DATE: 1998-01-05
; EARLIER APPLICATION NUMBER: PCT/SE96/00570
; EARLIER FILING DATE: 1996-05-02
; EARLIER APPLICATION NUMBER: SE 9501682-0
; EARLIER FILING DATE: 1995-05-05
; SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 5
LENGTH: 102
TYPE: PRT
ORGANISM: Vibrio cholerae
US-08-952-337-5
```

Query Match 100.0%; Score 62; DB 3; Length 102;
Best Local Similarity 100.0%; Pred. No. 0.00036;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 VEPGSOHDSQ 12
DB 49 VEPGSOHDSQ 60

RESULT 7
US-08-952-337-6
; Sequence 6, Application US/08952337
; Patent No. 601973
; GENERAL INFORMATION:
; APPLICANT: Holmgren, Jan
; APPLICANT: Lebens, Michael R.
; TITLE OF INVENTION: HYBRID MOLECULES BETWEEN HEAT-LABILE
; TITLE OF INVENTION: ENTEROTOXIN AND CHOLERA TOXIN B SUBUNITS
; FILE REFERENCE: 3846/00758
; CURRENT APPLICATION NUMBER: US/08/952,337
; CURRENT FILING DATE: 1998-01-05
; EARLIER APPLICATION NUMBER: PCT/SE96/00570
; EARLIER FILING DATE: 1996-05-02
; EARLIER APPLICATION NUMBER: SE 9501682-0
; EARLIER FILING DATE: 1995-05-05
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FASTSEQ for Windows Version 3.0
; SEQ ID NO 6
; LENGTH: 102
; TYPE: PRT
; ORGANISM: Escherichia coli
US-08-952-337-6

Query Match 100.0%; Score 62; DB 3; Length 102;
Best Local Similarity 100.0%; Pred. No. 0.00036;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 VEPGSOHDSQ 12
DB 49 VEPGSOHDSQ 60

RESULT 8
US-08-472-171-2
; Sequence 2, Application US/08472171
; Patent No. 5932714
; GENERAL INFORMATION:
; APPLICANT: Loomore, Sheena M.
; APPLICANT: Yacoub, Reza K.
; APPLICANT: Zealey, Gavin R.
; APPLICANT: Klein, Michel H.
; TITLE OF INVENTION: Expression of Gene Products From
; TITLE OF INVENTION: Genetically Manipulated Strains of Bordetella
; NUMBER OF SEQUENCES: 56
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: 330 University Avenue, Suite 701
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/472,171
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/393,334

; FILING DATE: 23-FEB-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Stewart, Michael I.
; REGISTRATION NUMBER: 24,973
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 416-595-1155
; TELEFAX: 416-595-1163
; TELEX: 065-24567 Simbas
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 103 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-472-171-2

Query Match 100.0%; Score 62; DB 2; Length 103;
Best Local Similarity 100.0%; Pred. No. 0.00036;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 VEPGSOHDSQ 12
DB 50 VEPGSOHDSQ 61

RESULT 9
US-08-894-526-2
; Sequence 2, Application US/08894526
; Patent No. 5942418
; GENERAL INFORMATION:
; APPLICANT: Loomore, Sheena M
; APPLICANT: Yacoub, Reza K
; APPLICANT: Zealey, Gavin R
; APPLICANT: Klein, Michel H
; TITLE OF INVENTION: EXPRESSION OF GENE PRODUCTS FROM
; TITLE OF INVENTION: GENETICALLY MANIPULATED STRAINS OF BORDETELLA
; NUMBER OF SEQUENCES: 57
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: 6th Floor, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/894,526
; FILING DATE: 01-DEC-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Stewart, Michael I
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-724 MIS-JB
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 103 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-894-526-2

Query Match 100.0%; Score 62; DB 2; Length 103;
Best Local Similarity 100.0%; Pred. No. 0.00036;

Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VEVPGSOHDSQ 12
|||||

Db 50 VEVPGSOHDSQ 61

RESULT 10

US-09-013-047-2

; Sequence 2, Application US/09013047
; Patent No. 5998168

; GENERAL INFORMATION:

; APPLICANT: Loosmore, Sheena M.

; APPLICANT: Yacoub, Reza K.

; APPLICANT: Zealey, Gavin R.

; APPLICANT: Klein, Michel H.

; TITLE OF INVENTION: Expression Of Gene Products From

; TITLE OF INVENTION: Genetically Manipulated Strains Of Bordetella

; NUMBER OF SEQUENCES: 56

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Sim & McBurney

; STREET: 330 University Avenue, 6th Floor

; CITY: Toronto

; STATE: Ontario

; COUNTRY: Canada

; ZIP: M5G 1R7

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/013,047

; FILING DATE:

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/472,171

; FILING DATE: 07-JUN-1995

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/393,334

; FILING DATE: 23-FEB-1995

; ATTORNEY/AGENT INFORMATION:

; NAME: Stewart, Michael I.

; REGISTRATION NUMBER: 24,973

; REFERENCE/DOCKET NUMBER: 1038-507 MIS:Vg

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 416-595-1155

; TELEFAX: 416-595-1163

; TELEX: 065-24567 Simbas

; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 103 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; US-09-013-047-2

Query Match 100.0%; Score 62; DB 2; Length 103;

Best Local Similarity 100.0%; Pred. No. 0.00036;

Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VEVPGSOHDSQ 12
|||||

Db 50 VEVPGSOHDSQ 61

RESULT 11

US-09-374-597-2

; Sequence 2, Application US/09374597

; Patent No. 6140082

; GENERAL INFORMATION:

; APPLICANT: Loosmore, Sheena M.

; APPLICANT: Yacoub, Reza K.

; APPLICANT: Zealey, Gavin R.

; APPLICANT: Klein, Michel H.

; TITLE OF INVENTION: Expression Of Gene Products From

; TITLE OF INVENTION: Genetically Manipulated Strains Of Bordetella

; NUMBER OF SEQUENCES: 56

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Sim & McBurney

; STREET: 330 University Avenue, 6th Floor

; CITY: Toronto

; STATE: Ontario

; COUNTRY: Canada

; ZIP: M5G 1R7

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/374,597

; FILING DATE:

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/393,334

; FILING DATE: FEBRUARY 23, 1995

; CLASSIFICATION:

; ATTORNEY/AGENT INFORMATION:

; NAME: Stewart, Michael I.

; REGISTRATION NUMBER: 24973

; REFERENCE/DOCKET NUMBER: 1038-964

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 416-595-1155

; TELEFAX: 416-595-1163

; TELEX: 065-24567 Simbas

; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 103 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; US-09-374-597-2

Query Match 100.0%; Score 62; DB 4; Length 103;

Best Local Similarity 100.0%; Pred. No. 0.00036;

Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VEVPGSOHDSQ 12
|||||

Db 50 VEVPGSOHDSQ 61

RESULT 12

US-09-191-852-21

; Sequence 21, Application US/09191852

; Patent No. 6194560

; GENERAL INFORMATION:

; APPLICANT: Charles J. Arentzen, Hugh S. Mason, and Tariq A. Haq

; TITLE OF INVENTION: ORAL IMMUNIZATION WITH TRANSGENIC PLANTS

; NUMBER OF SEQUENCES: 23

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Fulbright & Jaworski L.L.P.

; STREET: 1301 McKinnney, Suite 5100

; CITY: Houston

; STATE: TX

; COUNTRY: USA

; ZIP: 77010

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/191,852
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/13376
FILING DATE: 24-OCT-1995
APPLICATION NUMBER: 08/817,906
FILING DATE: 04-AUG-1997
ATTORNEY/AGENT INFORMATION:
NAME: Fox, David L.
REGISTRATION NUMBER: 40,612
REFERENCE/DOCKET NUMBER: P015900S1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 713-651-5151
TELEFAX: 713-651-5246
INFORMATION FOR SEQ ID NO: 21:
SEQUENCE CHARACTERISTICS:
LENGTH: 103 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-09-191-852-21

Query Match 100.0%; Score 62; DB 4; Length 103;
Best Local Similarity 100.0%; Pred. No. 0.00036;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VEVPGSQHDSQ 12
|||||
DB 50 VEVPGSQHDSQ 61

RESULT 13
PCT-US95-13376-21
Sequence 21, Application PC/TUS9513376
GENERAL INFORMATION:
APPLICANT: The Texas A&M University System
APPLICANT: 310 Wisenbaker
APPLICANT: College Station, Texas 77843-3369
TITLE OF INVENTION: ORAL IMMUNIZATION WITH TRANSGENIC PLANTS
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pravel, Hewitt, Kimball & Krieger
STREET: 1177 West Loop South, 10th Floor
CITY: Houston
STATE: TX
COUNTRY: USA
ZIP: 77027-9095
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/13376
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/328,716
FILING DATE: 24-OCT-1994
ATTORNEY/AGENT INFORMATION:
NAME: Jones, John W.
REGISTRATION NUMBER: 31,380
REFERENCE/DOCKET NUMBER: 36170/3P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 713-850-0909
TELEFAX: 713-850-0165
INFORMATION FOR SEQ ID NO: 21:
SEQUENCE CHARACTERISTICS:
LENGTH: 103 amino acids
TYPE: amino acid
TOPOLOGY: linear
PCT-US95-13376-21

Query Match 100.0%; Score 62; DB 5; Length 103;
Best Local Similarity 100.0%; Pred. No. 0.00036;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VEVPGSQHDSQ 12
|||||
DB 50 VEVPGSQHDSQ 61

RESULT 14
US-08-952-337-1
Sequence 1, Application US/08952337
Patent No. 6019973
GENERAL INFORMATION:
APPLICANT: Holmgren, Jan
APPLICANT: Lebens, Michael R.
TITLE OF INVENTION: HYBRID MOLECULES BETWEEN HEAT-LABILE
TITLE OF INVENTION: ENTEROTOXIN AND CHOLERA TOXIN B SUBUNITS
FILE REFERENCE: 3846/0D758
CURRENT APPLICATION NUMBER: US/08/952,337
EARLIER FILING DATE: 1998-01-05
EARLIER APPLICATION NUMBER: PCT/SE96/00570
EARLIER FILING DATE: 1996-05-02
EARLIER APPLICATION NUMBER: SE 9501682-0
EARLIER FILING DATE: 1995-05-05
NUMBER OF SEQ ID NOS: 6
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 1
LENGTH: 123
TYPE: PRT
ORGANISM: Vibrio cholerae
US-08-952-337-1

Query Match 100.0%; Score 62; DB 3; Length 123;
Best Local Similarity 100.0%; Pred. No. 0.00044;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VEVPGSQHDSQ 12
|||||
DB 70 VEVPGSQHDSQ 81

RESULT 15
US-08-952-337-2
Sequence 2, Application US/08952337
Patent No. 6019973
GENERAL INFORMATION:
APPLICANT: Holmgren, Jan
APPLICANT: Lebens, Michael R.
TITLE OF INVENTION: ENTEROTOXIN AND CHOLERA TOXIN B SUBUNITS
FILE REFERENCE: 3846/0D758
CURRENT APPLICATION NUMBER: US/08/952,337
EARLIER FILING DATE: 1998-01-05
EARLIER APPLICATION NUMBER: PCT/SE96/00570
EARLIER FILING DATE: 1996-05-02
EARLIER APPLICATION NUMBER: SE 9501682-0
EARLIER FILING DATE: 1995-05-05
NUMBER OF SEQ ID NOS: 6
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 2
LENGTH: 123
TYPE: PRT
ORGANISM: Escherichia coli
US-08-952-337-2

Query Match 100.0%; Score 62; DB 3; Length 123;
Best Local Similarity 100.0%; Pred. No. 0.00044;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Mon Jul 8 07:51:48 2002

us-09-786-648-3.rai

Page 7

Oy 1 VEVPGSQHDSQ 12
|||||||
Db 70 VEVPGSQHDSQ 81

Search completed: July 3, 2002, 08:55:39
Job time: 214 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 3, 2002, 09:03:20 ; Search time 401.04 seconds
(without alignments)
10.532 Million cell updates/sec

Title: US-09-786-648-3

Perfect score: 62

Sequence: 1 VEPGSHDSQ 12

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 3502263 seqs, 351980561 residues

Total number of hits satisfying chosen parameters: 3502263

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%.

Listing First 45 summaries

Database : Pending_Patents_AA_Main:*

1:	/cgn2_6/ptodata/2/paa/US086_COMB.pep.*	Sequence 26, Appl
2:	/cgn2_6/ptodata/2/paa/US086_COMB.pep.*	Sequence 2, Appl
3:	/cgn2_6/ptodata/2/paa/US086_COMB.pep.*	Sequence 15, Appl
4:	/cgn2_6/ptodata/2/paa/US086_COMB.pep.*	Sequence 21, Appl
5:	/cgn2_6/ptodata/2/paa/US086_COMB.pep.*	Sequence 10, Appl
6:	/cgn2_6/ptodata/2/paa/US086_COMB.pep.*	Sequence 22, Appl
7:	/cgn2_6/ptodata/2/paa/US086_COMB.pep.*	Sequence 52, Appl
8:	/cgn2_6/ptodata/2/paa/US086_COMB.pep.*	Sequence 57, Appl
9:	/cgn2_6/ptodata/2/paa/US086_COMB.pep.*	Sequence 55, Appl
10:	/cgn2_6/ptodata/2/paa/US086_COMB.pep.*	Sequence 57, Appl
11:	/cgn2_6/ptodata/2/paa/US086_COMB.pep.*	Sequence 2, Appl
12:	/cgn2_6/ptodata/2/paa/US086_COMB.pep.*	Sequence 2, Appl
13:	/cgn2_6/ptodata/2/paa/US086_COMB.pep.*	Sequence 15, Appl
14:	/cgn2_6/ptodata/2/paa/US086_COMB.pep.*	Sequence 2, Appl
15:	/cgn2_6/ptodata/2/paa/US086_COMB.pep.*	Sequence 2, Appl
16:	/cgn2_6/ptodata/2/paa/US086_COMB.pep.*	Sequence 2, Appl
17:	/cgn2_6/ptodata/2/paa/US086_COMB.pep.*	Sequence 2, Appl
18:	/cgn2_6/ptodata/2/paa/US086_COMB.pep.*	Sequence 2, Appl
19:	/cgn2_6/ptodata/2/paa/US086_COMB.pep.*	Sequence 2, Appl
20:	/cgn2_6/ptodata/2/paa/US086_COMB.pep.*	Sequence 2, Appl
21:	/cgn2_6/ptodata/2/paa/US086_COMB.pep.*	Sequence 2, Appl
22:	/cgn2_6/ptodata/2/paa/US086_COMB.pep.*	Sequence 2, Appl
23:	/cgn2_6/ptodata/2/paa/US086_COMB.pep.*	Sequence 2, Appl
24:	/cgn2_6/ptodata/2/paa/US086_COMB.pep.*	Sequence 2, Appl
25:	/cgn2_6/ptodata/2/paa/US086_COMB.pep.*	Sequence 2, Appl
26:	/cgn2_6/ptodata/2/paa/US086_COMB.pep.*	Sequence 2, Appl

ALIGNMENTS

RESULT 1

US-09-786-648-3

Sequence 3, Application US/09786648

GENERAL INFORMATION:

APPLICANT: Williams, Neil Andrew

TITLE OR INVENTION: Peptide Fragments of Cholera Toxin B or Enterotoxin B as Vacc

FILE REFERENCE: 7438

CURRENT APPLICATION NUMBER: US/09/786, 648

CURRENT FILING DATE: 2001-03-07

PRIOR APPLICATION NUMBER: PCT/GB99/02970

PRIOR FILING DATE: 1999-09-07

NUMBER OF SEQ ID NOS: 6

SOFTWARE: MS DOS

SEQ ID NO 3

LENGTH: 12

TYPE: PRT

ORGANISM: E. coli

FEATURE:

LOCATION: 50...61

OTHER INFORMATION: isolated or synthetic EtXB beta4-alpha2 loop fragment derivabl

OTHER INFORMATION: human variant E. coli

US-09-786-648-3

Result No.	Score	Query Match	Length DB	ID	Description
1	62	100.0	12	21	US-09-786-648-3
2	62	100.0	15	11	US-08-732-371-1
3	62	100.0	15	11	US-08-732-371A-1
4	62	100.0	21	21	US-09-786-648-4
5	62	100.0	21	21	US-09-786-648-5
6	62	100.0	93	5	US-08-110-947-10
7	62	100.0	93	5	US-08-110-947A-26

SUMMARIES

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Query Match 100.0%; Score 62; DB 21; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.00041;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VEPGSHIDSQ 12
|||||
DB 1 VEPGSHIDSQ 12

RESULT 2

US-08-732-371-1
; Sequence 1, Application US/08732371
; GENERAL INFORMATION:
; APPLICANT: MIRELMAN, David
; APPLICANT: MARKS, Robert S.
; APPLICANT: SELA, Michael
; TITLE OF INVENTION: VACCINES FOR ORAL IMMUNIZATION AGAINST
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWDY AND NEIMARK
; STREET: 419 Seventh Street, N.W., Suite 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/732,371
; FILING DATE: 09-JAN-1997
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: IL 109519
; FILING DATE: 03-MAY-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: YUN, Allen C.
; REGISTRATION NUMBER: 37,971
; REFERENCE/DOCKET NUMBER: MIRELMAN-3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-628-5197
; TELEFAX: 202-737-3528
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-732-371-1

Query Match 100.0%; Score 62; DB 11; Length 15;
Best Local Similarity 100.0%; Pred. No. 0.00053;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 VEPGSHIDSQ 12
|||||
DB 1 VEPGSHIDSQ 12

RESULT 3

US-08-732-371A-1
; Sequence 1, Application US/08732371A
; GENERAL INFORMATION:
; APPLICANT: MIRELMAN, David
; APPLICANT: MARKS, Robert S.
; APPLICANT: SELA, Michael
; TITLE OF INVENTION: VACCINES FOR ORAL IMMUNIZATION AGAINST
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWDY AND NEIMARK
; STREET: 419 Seventh Street, N.W., Suite 300
; CITY: Washington

STATE: D.C.
COUNTRY: USA
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/732,371A
FILING DATE: 09-JAN-1997
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: IL 109519
FILING DATE: 03-MAY-1994
ATTORNEY/AGENT INFORMATION:
NAME: YUN, Allen C.
REGISTRATION NUMBER: 37,971
REFERENCE/DOCKET NUMBER: MIRELMAN-3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-628-5197
TELEFAX: 202-737-3528
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 15 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-732-371A-1

Query Match 100.0%; Score 62; DB 11; Length 15;
Best Local Similarity 100.0%; Pred. No. 0.00053;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VEPGSHIDSQ 12
|||||
DB 1 VEPGSHIDSQ 12

RESULT 4
US-09-786-648-4
; Sequence 4, Application US/09786648
; GENERAL INFORMATION:
; APPLICANT: Williams, Neil Andrew
; APPLICANT: Hirst, Timothy Raymond
; TITLE OF INVENTION: Peptide Fragments of Cholera Toxin B or Enterotoxin B as Vacci
; FILE REFERENCE: 7438
; CURRENT APPLICATION NUMBER: US/09/786,648
; CURRENT FILING DATE: 2001-03-07
; PRIOR APPLICATION NUMBER: PCT/GB99/02970
; PRIOR FILING DATE: 1999-09-07
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: MS DOS
; SEQ ID NO: 4
; LENGTH: 21
; TYPE: PRT
; ORGANISM: E. coli
; FEATURE:
; LOCATION: 45...65
; OTHER INFORMATION: Isolated or synthetic EtXB beta4-alpha2 loop fragment derivabl
; OTHER INFORMATION: human variant E. coli
US-09-786-648-4

Query Match 100.0%; Score 62; DB 21; Length 21;
Best Local Similarity 100.0%; Pred. No. 0.00079;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VEPGSHIDSQ 12
|||||
DB 6 VEPGSHIDSQ 17

RESULT 5
US-09-786-648-5
; Sequence 5, Application US/09786648
; GENERAL INFORMATION:
; APPLICANT: Williams, Neil Andrew
; APPLICANT: Hirst, Timothy Raymond
; TITLE OF INVENTION: Peptide Fragments of Cholera Toxin B or Enterotoxin B as Vaccine
; FILE REFERENCE: 7438
; CURRENT APPLICATION NUMBER: US/09/786,648
; CURRENT FILING DATE: 2001-03-07
; PRIOR APPLICATION NUMBER: PCT/GB99/02970
; PRIOR FILING DATE: 1999-09-07
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: MS DOS
; SEQ ID NO 5
; LENGTH: 21
; TYPE: PRT
; ORGANISM: E. coli
; FEATURE:
; LOCATION: 45...65
; OTHER INFORMATION: Isolated or synthetic ElxB beta4-alpha2 loop fragment derivable f
; OTHER INFORMATION: porcine E. coli
US-09-786-648-5

Query Match 100.0%; Score 62; DB 21; Length 21;
Best Local Similarity 100.0%; Pred. No. 0.00079;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VEPGSOHDSQ 12
|||||
DB 6 VEPGSOHDSQ 17

RESULT 6
US-08-110-947-10
; Sequence 10, Application US/08110947
; GENERAL INFORMATION:
; APPLICANT: READ, Randy J
; APPLICANT: STEIN, Penelope E
; APPLICANT: COCKLE, Stephen A
; APPLICANT: COHEN, Raymond P
; APPLICANT: KLEIN, Michel H
; TITLE OF INVENTION: MODIFICATION OF PERTUSSIS TOXIN
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Shoemaker and Mattare, Ltd.
; STREET: Suite 1203 Crystal Plaza I, 2001 Jefferson
; STREET: Davis Hwy.
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202-0286
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/110,947
; FILING DATE: 24-AUG-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: FALLON, Charles W
; REGISTRATION NUMBER: 28,946
; REFERENCE/DOCKET NUMBER: 1038-303 MIS:jb
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 415-0810
; TELEFAX: (703) 415-0813
; TELEFAX: 89-9456 LUKPAT

; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 93 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-110-947-10

Query Match 100.0%; Score 62; DB 5; Length 93;
Best Local Similarity 100.0%; Pred. No. 0.0044;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VEPGSOHDSQ 12
|||||
DB 40 VEPGSOHDSQ 51

RESULT 7
US-08-110-947A-26
; Sequence 26, Application US/08110947A
; GENERAL INFORMATION:
; APPLICANT: READ, Randy J
; APPLICANT: STEIN, Penelope E
; APPLICANT: COCKLE, Stephen A
; APPLICANT: COHEN, Raymond P
; APPLICANT: KLEIN, Michel H
; TITLE OF INVENTION: MODIFICATION OF PERTUSSIS TOXIN
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Shoemaker and Mattare, Ltd.
; STREET: Suite 1203 Crystal Plaza I, 2001 Jefferson
; STREET: Davis Highway
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202-0286
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/110,947A
; FILING DATE: 24-AUG-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: FALLON, Charles W
; REGISTRATION NUMBER: 28,946
; REFERENCE/DOCKET NUMBER: 1038-303 MIS:jb
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 415-0810
; TELEFAX: (703) 415-0813
; TELEFAX: 89-9456 LUKPAT
; INFORMATION FOR SEQ ID NO: 26:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 93 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-110-947A-26

Query Match 100.0%; Score 62; DB 5; Length 93;
Best Local Similarity 100.0%; Pred. No. 0.0044;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VEPGSOHDSQ 12
|||||
DB 40 VEPGSOHDSQ 51

RESULT 8

US-08-251-121-26
; Sequence 26, Application US/08251121
; GENERAL INFORMATION:
; APPLICANT: READ, Randy J.
; APPLICANT: STEIN, Penelope E.
; APPLICANT: COCKLE, Stephen A.
; APPLICANT: OOMEN, Raymond P.
; APPLICANT: KLEIN, Michel H.
; APPLICANT: ARMSTRONG, Glen D.
; APPLICANT: HAZES, Bart
; TITLE OF INVENTION: MODIFICATION OF PERTUSSIS TOXIN
; NUMBER OF SEQUENCES: 46
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Slim & McBurney
; STREET: Suite 701, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/251,121
; FILING DATE: 31-MAY-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/110,947
; FILING DATE: 24-AUG-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: STEWART, Michael I.
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-335
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; INFORMATION FOR SEQ ID NO: 26:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 93 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-251-121-26

Query Match 100.0%; Score 62; DB 6; Length 93;
Best Local Similarity 100.0%; Pred. No. 0.0044;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VEVPSQHIHSQ 12
DB 40 VEVPSQHIHSQ 51

RESULT 9
US-08-393-334-2
; Sequence 2, Application US/08393334
; GENERAL INFORMATION:
; APPLICANT: LOOSMORE, Sheena M.
; APPLICANT: YACCOB, Reza K.
; APPLICANT: ZEALY, Gavin R.
; APPLICANT: KLEIN, Michel H.
; TITLE OF INVENTION: Expression Of Gene Products From
; TITLE OF INVENTION: Genetically Manipulated Strains Of Bordetella
; NUMBER OF SEQUENCES: 56
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Slim & McBurney
; STREET: 330 University Avenue, Suite 701
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada

ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/393,334
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Stewart, Michael I.
; REGISTRATION NUMBER: 24973
; REFERENCE/DOCKET NUMBER: 1038-417
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 416-595-1155
; TELEFAX: 416-595-1163
; TELEX: 065-24567 Simbas
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 103 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-393-334-2

Query Match 100.0%; Score 62; DB 7; Length 103;
Best Local Similarity 100.0%; Pred. No. 0.005;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VEVPSQHIHSQ 12
DB 50 VEVPSQHIHSQ 61

RESULT 10
US-08-782-832-15
; Sequence 15, Application US/08782832
; GENERAL INFORMATION:
; APPLICANT: ARNTZEN, Charles J.
; APPLICANT: HAG, Tariq A.
; TITLE OF INVENTION: PRODUCTION OF AN ORALLY IMMUNOGENIC
; TITLE OF INVENTION: BACTERIAL PROTEIN IN TRANSGENIC PLANTS
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pravel, Hewitt, Kimball & Krieger
; STREET: 1477 West Loop South, 10th Floor
; CITY: Houston
; STATE: TX
; COUNTRY: USA
; ZIP: 77027-9095
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/782,832
; FILING DATE: 13-JAN-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/328,716
; FILING DATE: 24-OCT-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Kettelberger, Denise M.
; REGISTRATION NUMBER: 33,924
; REFERENCE/DOCKET NUMBER: 36170/3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 713-850-0909
; TELEFAX: 713-850-0165
; INFORMATION FOR SEQ ID NO: 15:

SEQUENCE CHARACTERISTICS:
 LENGTH: 103 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 US-08-782-832-15

Query Match 100.0%; Score 62; DB 11; Length 103;
 Best Local Similarity 100.0%; Pred. No. 0.005;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VEVPGSQHDSQ 12
 |||||||||
 DB 50 VEVPGSQHDSQ 61

RESULT 11
 US-08-817-906-21

; Sequence 21, Application US/08817906
 ; GENERAL INFORMATION:
 ; APPLICANT: Charles J. Arntzen, Hugh S. Mason, John D. Clements,
 ; APPLICANT: and Tariq A. Haq
 ; TITLE OF INVENTION: ORAL IMMUNIZATION WITH TRANSGENIC PLANTS
 ; NUMBER OF SEQUENCES: 23
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Fulbright & Jaworski L.L.P.
 ; STREET: 1301 McKinney, Suite 5100
 ; CITY: Houston
 ; STATE: TX
 ; COUNTRY: USA
 ; ZIP: 77010
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/817,906
 ; FILING DATE: 08/04/97
 ; CLASSIFICATION: 800
 ; PRIORITY APPLICATION DATA:
 ; APPLICATION NUMBER: PCT/US95/13376
 ; FILING DATE: 24-OCT-1995
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Fox, David L.
 ; REGISTRATION NUMBER: 40,612
 ; REFERENCE/DOCKET NUMBER: 36170/3P
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 713-651-5151
 ; TELEFAX: 713-651-5246
 ; INFORMATION FOR SEQ ID NO: 21:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 103 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; US-08-817-906-21

Query Match 100.0%; Score 62; DB 12; Length 103;
 Best Local Similarity 100.0%; Pred. No. 0.005;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VEVPGSQHDSQ 12
 |||||||||
 DB 50 VEVPGSQHDSQ 61

RESULT 12
 US-09-836-433-14

; Sequence 14, Application US/09836433
 ; GENERAL INFORMATION:
 ; APPLICANT: Yuki, Yoshikazu
 ; APPLICANT: Uda, Shigezo

; TITLE OF INVENTION: HYBRID PROTEINS FOR AUTOIMMUNE DISEASE
 ; FILE REFERENCE: Not Assigned
 ; CURRENT APPLICATION NUMBER: US/09/836,433
 ; CURRENT FILING DATE: 2001-04-16
 ; NUMBER OF SEQ ID NOS: 49
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 14
 ; LENGTH: 103
 ; TYPE: PRT
 ; ORGANISM: Vibrio cholerae
 ; US-09-836-433-14

Query Match 100.0%; Score 62; DB 22; Length 103;
 Best Local Similarity 100.0%; Pred. No. 0.005;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VEVPGSQHDSQ 12
 |||||||||
 DB 50 VEVPGSQHDSQ 61

RESULT 13
 US-09-836-433-20

; Sequence 20, Application US/09836433
 ; GENERAL INFORMATION:
 ; APPLICANT: Yuki, Yoshikazu
 ; APPLICANT: Uda, Shigezo
 ; TITLE OF INVENTION: HYBRID PROTEINS FOR AUTOIMMUNE DISEASE
 ; FILE REFERENCE: Not Assigned
 ; CURRENT APPLICATION NUMBER: US/09/836,433
 ; CURRENT FILING DATE: 2001-04-16
 ; NUMBER OF SEQ ID NOS: 49
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 20
 ; LENGTH: 116
 ; TYPE: PRT
 ; ORGANISM: synthetic construct
 ; US-09-836-433-20

Query Match 100.0%; Score 62; DB 22; Length 116;
 Best Local Similarity 100.0%; Pred. No. 0.005;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VEVPGSQHDSQ 12
 |||||||||
 DB 50 VEVPGSQHDSQ 61

RESULT 14
 US-09-836-433-22

; Sequence 22, Application US/09836433
 ; GENERAL INFORMATION:
 ; APPLICANT: Yuki, Yoshikazu
 ; APPLICANT: Uda, Shigezo
 ; TITLE OF INVENTION: HYBRID PROTEINS FOR AUTOIMMUNE DISEASE
 ; FILE REFERENCE: Not Assigned
 ; CURRENT APPLICATION NUMBER: US/09/836,433
 ; CURRENT FILING DATE: 2001-04-16
 ; NUMBER OF SEQ ID NOS: 49
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 22
 ; LENGTH: 119
 ; TYPE: PRT
 ; ORGANISM: synthetic construct
 ; US-09-836-433-22

Query Match 100.0%; Score 62; DB 22; Length 119;
 Best Local Similarity 100.0%; Pred. No. 0.0059;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VEVPGSQHDSQ 12
 |||||
 Db 50 VEVPGSQHDSQ 61

RESULT 15
 PCT-US99-30747-55
 ; Sequence 55, Application PC/TUS9930747
 ; GENERAL INFORMATION:
 ; APPLICANT: Boyce Thompson Institute for Plant Research at Cor
 ; TITLE OF INVENTION: Orally Immunogenic Bacterial Enterotoxins Expressed In
 ; TITLE OF INVENTION: Transgenic Plants
 ; FILE REFERENCE: 4868/85427
 ; CURRENT APPLICATION NUMBER: PCT/US99/30747
 ; CURRENT FILING DATE: 1999-12-22
 ; NUMBER OF SEQ ID NOS: 67
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 55
 ; LENGTH: 124
 ; TYPE: PRT
 ; ORGANISM: Escherichia coli
 PCT-US99-30747-55

Query Match 100.0%; Score 62; DB 1; Length 124;
 Best Local Similarity 100.0%; Pred. No. 0.0062;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VEVPGSQHDSQ 12
 |||||
 Db 71 VEVPGSQHDSQ 82

Search completed: July 3, 2002, 09:03:20
 Job time: 620 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 3, 2002, 09:04:12 ; Search time 46.52 Seconds
(without alignments)
26.099 Million cell updates/sec

Title: US-09-786-648-3

Perfect score: 62

Sequence: 1 VEPVGSQHDSQ 12

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 308740 seqs, 101176262 residues

Total number of hits satisfying chosen parameters: 308740

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending Patents AA.New:*

1: /cgn2_6/ptodata/2/paa/PCT_NEW_COMB.pep:*
2: /cgn2_6/ptodata/2/paa/US06_NEW_COMB.pep:*
3: /cgn2_6/ptodata/2/paa/US07_NEW_COMB.pep:*
4: /cgn2_6/ptodata/2/paa/US08_NEW_COMB.pep:*
5: /cgn2_6/ptodata/2/paa/US09_NEW_COMB.pep:*
6: /cgn2_6/ptodata/2/paa/US10_NEW_COMB.pep:*
7: /cgn2_6/ptodata/2/paa/US60_NEW_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	62	100.0	103	6	US-10-110-364-8
2	62	100.0	103	6	US-10-110-364-10
3	62	100.0	103	6	US-10-110-364-13
4	62	100.0	103	6	US-10-110-364-12
5	62	100.0	105	6	US-10-110-364-22
6	62	100.0	123	6	US-10-110-364-17
7	62	100.0	123	6	US-10-110-364-20
8	62	100.0	124	6	US-10-110-364-2
9	62	100.0	124	6	US-10-110-364-6
10	62	100.0	124	6	US-10-110-364-11
11	62	100.0	124	6	US-10-110-364-16
12	62	100.0	124	6	US-10-110-364-18
13	62	100.0	124	6	US-10-110-364-19
14	62	100.0	124	6	US-10-110-364-21
15	62	100.0	124	6	US-10-110-364-23
16	62	100.0	125	6	US-10-110-364-15
17	62	100.0	138	6	US-10-141-627-2
18	56	90.3	103	6	US-10-110-364-5
19	56	90.3	103	6	US-10-110-364-7
20	56	90.3	124	6	US-10-110-364-4
21	56	90.3	124	6	US-10-110-364-9
22	56	90.3	382	5	US-09-809-033A-3
23	56	90.3	382	5	US-09-809-033A-4
24	42	67.7	2359	6	US-10-155-881-28520
25	39	62.9	412	5	US-09-978-403A-157
26	39	62.9	412	5	US-09-978-544A-157

27	39	62.9	412	5	US-09-978-681A-157	Sequence 157, App
28	39	62.9	412	5	US-09-978-757A-157	Sequence 157, App
29	39	62.9	412	5	US-09-978-564A-157	Sequence 157, App
30	39	62.9	412	5	US-09-999-831A-157	Sequence 157, App
31	39	62.9	412	5	US-09-999-829A-157	Sequence 157, App
32	39	62.9	412	5	US-09-978-375A-157	Sequence 157, App
33	39	62.9	412	5	US-09-978-423A-157	Sequence 157, App
34	39	62.9	412	6	US-10-013-921A-157	Sequence 157, App
35	39	62.9	412	6	US-10-013-929A-157	Sequence 157, App
36	39	62.9	412	6	US-10-013-918A-157	Sequence 157, App
37	39	62.9	412	6	US-10-017-082A-157	Sequence 157, App
38	39	62.9	412	6	US-10-017-085A-157	Sequence 157, App
39	39	62.9	412	6	US-10-013-916A-157	Sequence 157, App
40	39	62.9	412	6	US-10-017-086A-157	Sequence 157, App
41	39	62.9	412	6	US-10-013-925A-157	Sequence 157, App
42	39	62.9	412	6	US-10-017-081A-157	Sequence 157, App
43	39	62.9	412	6	US-10-016-177A-157	Sequence 157, App
44	39	62.9	412	6	US-10-017-084A-157	Sequence 157, App
45	39	62.9	412	6	US-10-013-923A-157	Sequence 157, App

ALIGNMENTS

```
RESULT 1
; Sequence 8, Application US/10110364
; GENERAL INFORMATION:
; APPLICANT: Handley, Harold H.
; APPLICANT: Haaparanta, Tapio
; APPLICANT: Ewall, Karla L.
; TITLE OF INVENTION: ABS TOXIN B SUBUNIT MUTANTS WITH ALTERED
; FILE REFERENCE: ACTBIO.004A
; CURRENT APPLICATION NUMBER: US/10/110,364
; PRIOR FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: PCT/US00/27607
; PRIOR FILING DATE: 2000-10-05
; PRIOR APPLICATION NUMBER: 60/158,561
; PRIOR FILING DATE: 1999-10-08
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 8
; LENGTH: 103
; TYPE: PRT
; ORGANISM: Vibrio cholera
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(103)
; OTHER INFORMATION: CTB variant from NCBI gene bank GI: 998409
US-10-110-364-8

Query Match          100.0%  Score 62:  DB 6:  Length 103:
Best Local Similarity 100.0%  Pred No. 0.00021:
Matches 12:  Conservative 0:  Mismatches 0:  Indels 0:  Gaps 0:

QY      1 VEPVGSQHDSQ 12
Db      50 VEPVGSQHDSQ 61

RESULT 2
; Sequence 10, Application US/10110364
; GENERAL INFORMATION:
; APPLICANT: Handley, Harold H.
; APPLICANT: Haaparanta, Tapio
; APPLICANT: Ewall, Karla L.
; TITLE OF INVENTION: ABS TOXIN B SUBUNIT MUTANTS WITH ALTERED
; FILE REFERENCE: ACTBIO.004A
; CURRENT APPLICATION NUMBER: US/10/110,364
```

```

; CURRENT FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: PCT/US00/27607
; PRIOR FILING DATE: 2000-10-05
; PRIOR APPLICATION NUMBER: 60/158,561
; PRIOR FILING DATE: 1999-10-08
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 103
; TYPE: PRT
; ORGANISM: Vibrio cholera
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(103)
; OTHER INFORMATION: CTB variant from NCBI gene bank GI: 1421511
US-10-110-364-10

Query Match          100.0%; Score 62; DB 6; Length 103;
Best Local Similarity 100.0%; Pred. No. 0.00021;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 VEPGSHIDSO 12
        |||||||
Db       50 VEPGSHIDSO 61

RESULT      3
US-10-110-364-13
; Sequence 13, Application US/10110364
; GENERAL INFORMATION:
; APPLICANT: Handley, Harold H.
; APPLICANT: Haaparanta, Tapio
; APPLICANT: Ewalt, Karla L.
; TITLE OF INVENTION: AB5 TOXIN B SUBUNIT MUTANTS WITH ALTERED
; TITLE OF INVENTION: CHEMICAL CONJUGATION CHARACTERISTICS
; FILE REFERENCE: ACTBIO.004A
; CURRENT APPLICATION NUMBER: US/10/110,364
; PRIOR FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: PCT/US00/27607
; PRIOR FILING DATE: 2000-10-05
; PRIOR APPLICATION NUMBER: 60/158,561
; PRIOR FILING DATE: 1999-10-08
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 103
; TYPE: PRT
; ORGANISM: Vibrio cholera
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(103)
; OTHER INFORMATION: CTB variant from NCBI gene bank GI: 14215235
US-10-110-364-13

Query Match          100.0%; Score 62; DB 6; Length 103;
Best Local Similarity 100.0%; Pred. No. 0.00021;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 VEPGSHIDSO 12
        |||||||
Db       50 VEPGSHIDSO 61

RESULT      4
US-10-110-364-22
; Sequence 22, Application US/10110364
; GENERAL INFORMATION:
; APPLICANT: Handley, Harold H.
; APPLICANT: Haaparanta, Tapio
; APPLICANT: Ewalt, Karla L.
```

```

; TITLE OF INVENTION: AB5 TOXIN B SUBUNIT MUTANTS WITH ALTERED
; TITLE OF INVENTION: CHEMICAL CONJUGATION CHARACTERISTICS
; FILE REFERENCE: ACTBIO.004A
; CURRENT APPLICATION NUMBER: US/10/110,364
; PRIOR FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: PCT/US00/27607
; PRIOR FILING DATE: 2000-10-05
; PRIOR APPLICATION NUMBER: 60/158,561
; PRIOR FILING DATE: 1999-10-08
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 22
; LENGTH: 103
; TYPE: PRT
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(103)
; OTHER INFORMATION: LTB variant from NCBI gene bank GI: 494265.
US-10-110-364-22

Query Match          100.0%; Score 62; DB 6; Length 103;
Best Local Similarity 100.0%; Pred. No. 0.00021;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 VEPGSHIDSO 12
        |||||||
Db       50 VEPGSHIDSO 61

RESULT      5
US-10-110-364-12
; Sequence 12, Application US/10110364
; GENERAL INFORMATION:
; APPLICANT: Handley, Harold H.
; APPLICANT: Haaparanta, Tapio
; APPLICANT: Ewalt, Karla L.
; TITLE OF INVENTION: AB5 TOXIN B SUBUNIT MUTANTS WITH ALTERED
; TITLE OF INVENTION: CHEMICAL CONJUGATION CHARACTERISTICS
; FILE REFERENCE: ACTBIO.004A
; CURRENT APPLICATION NUMBER: US/10/110,364
; PRIOR FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: PCT/US00/27607
; PRIOR FILING DATE: 2000-10-05
; PRIOR APPLICATION NUMBER: 60/158,561
; PRIOR FILING DATE: 1999-10-08
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 105
; TYPE: PRT
; ORGANISM: Vibrio cholera
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(105)
; OTHER INFORMATION: CTB variant from NCBI gene bank GI: 2781121 (Ogawa
US-10-110-364-12

Query Match          100.0%; Score 62; DB 6; Length 105;
Best Local Similarity 100.0%; Pred. No. 0.00021;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 VEPGSHIDSO 12
        |||||||
Db       51 VEPGSHIDSO 62

RESULT      6
US-10-110-364-17
; Sequence 17, Application US/10110364
```

```

; GENERAL INFORMATION:
; APPLICANT: Handley, Harold H.
; APPLICANT: Haaparanta, Tapio
; APPLICANT: Ewalt, Karla L.
; TITLE OF INVENTION: AB5 TOXIN B SUBUNIT MUTANTS WITH ALTERED
; TITLE OF INVENTION: CHEMICAL CONJUGATION CHARACTERISTICS
; FILE REFERENCE: ACTBIO.004A
; CURRENT APPLICATION NUMBER: US/10/110,364
; CURRENT FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: PCT/US00/27607
; PRIOR FILING DATE: 2000-10-05
; PRIOR APPLICATION NUMBER: 60/158,561
; PRIOR FILING DATE: 1999-10-08
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 123
; TYPE: PRT
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(123)
; OTHER INFORMATION: LTB variant from NCBI gene bank GI: 1395122.
US-10-110-364-17

```

```

Query Match      100.0%; Score 62; DB 6; Length 123;
Best Local Similarity 100.0%; Pred. No. 0.00025;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 VEPGSOHDSQ 12
   |||||
Db 71 VEPGSOHDSQ 82

```

```

RESULT 7
US-10-110-364-20
; Sequence 20, Application US/10110364
; GENERAL INFORMATION:
; APPLICANT: Handley, Harold H.
; APPLICANT: Haaparanta, Tapio
; APPLICANT: Ewalt, Karla L.
; TITLE OF INVENTION: AB5 TOXIN B SUBUNIT MUTANTS WITH ALTERED
; TITLE OF INVENTION: CHEMICAL CONJUGATION CHARACTERISTICS
; FILE REFERENCE: ACTBIO.004A
; CURRENT APPLICATION NUMBER: US/10/110,364
; CURRENT FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: PCT/US00/27607
; PRIOR FILING DATE: 2000-10-05
; PRIOR APPLICATION NUMBER: 60/158,561
; PRIOR FILING DATE: 1999-10-08
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 20
; LENGTH: 123
; TYPE: PRT
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(123)
; OTHER INFORMATION: LTB variant from NCBI gene bank GI: 223254.
US-10-110-364-20

```

```

Query Match      100.0%; Score 62; DB 6; Length 123;
Best Local Similarity 100.0%; Pred. No. 0.00025;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 VEPGSOHDSQ 12
   |||||
Db 70 VEPGSOHDSQ 81

```

```

RESULT 8
US-10-110-364-2
; Sequence 2, Application US/10110364
; GENERAL INFORMATION:
; APPLICANT: Handley, Harold H.
; APPLICANT: Haaparanta, Tapio
; APPLICANT: Ewalt, Karla L.
; TITLE OF INVENTION: AB5 TOXIN B SUBUNIT MUTANTS WITH ALTERED
; TITLE OF INVENTION: CHEMICAL CONJUGATION CHARACTERISTICS
; FILE REFERENCE: ACTBIO.004A
; CURRENT APPLICATION NUMBER: US/10/110,364
; CURRENT FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: PCT/US00/27607
; PRIOR FILING DATE: 2000-10-05
; PRIOR APPLICATION NUMBER: 60/158,561
; PRIOR FILING DATE: 1999-10-08
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 124
; TYPE: PRT
; ORGANISM: Vibrio cholera
US-10-110-364-2

```

```

Query Match      100.0%; Score 62; DB 6; Length 124;
Best Local Similarity 100.0%; Pred. No. 0.00026;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 VEPGSOHDSQ 12
   |||||
Db 71 VEPGSOHDSQ 82

```

```

RESULT 9
US-10-110-364-6
; Sequence 6, Application US/10110364
; GENERAL INFORMATION:
; APPLICANT: Handley, Harold H.
; APPLICANT: Haaparanta, Tapio
; APPLICANT: Ewalt, Karla L.
; TITLE OF INVENTION: AB5 TOXIN B SUBUNIT MUTANTS WITH ALTERED
; TITLE OF INVENTION: CHEMICAL CONJUGATION CHARACTERISTICS
; FILE REFERENCE: ACTBIO.004A
; CURRENT APPLICATION NUMBER: US/10/110,364
; CURRENT FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: PCT/US00/27607
; PRIOR FILING DATE: 2000-10-05
; PRIOR APPLICATION NUMBER: 60/158,561
; PRIOR FILING DATE: 1999-10-08
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 124
; TYPE: PRT
; ORGANISM: Vibrio cholera
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(124)
; OTHER INFORMATION: CTB variant from NCBI gene bank GI: 808900.
US-10-110-364-6

```

```

Query Match      100.0%; Score 62; DB 6; Length 124;
Best Local Similarity 100.0%; Pred. No. 0.00026;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 VEPGSOHDSQ 12
   |||||
Db 71 VEPGSOHDSQ 82

```

RESULT 10

US-10-110-364-11
; Sequence 11, Application US/10110364
; GENERAL INFORMATION:
; APPLICANT: Handley, Harold H.
; APPLICANT: Haaparanta, Tapio
; APPLICANT: Ewalt, Karla L.
; TITLE OF INVENTION: ABS TOXIN B SUBUNIT MUTANTS WITH ALTERED
; TITLE OF INVENTION: CHEMICAL CONJUGATION CHARACTERISTICS
; FILE REFERENCE: ACTBIO.004A
; CURRENT APPLICATION NUMBER: US/10/110,364
; CURRENT FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: PCT/US00/27607
; PRIOR FILING DATE: 2000-10-05
; PRIOR APPLICATION NUMBER: 60/158,561
; PRIOR FILING DATE: 1999-10-08
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 124
; TYPE: PRT
; ORGANISM: Vibrio cholera
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(124)
; OTHER INFORMATION: CTB variant from NCBI gene bank GI: 48890 (classic)
; OTHER INFORMATION: 569B).
US-10-110-364-11

Query Match 100.0%; Score 62; DB 6; Length 124;
Best Local Similarity 100.0%; Pred. No. 0.00026;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VEVPGSQHIDSQ 12
|||||
Db 71 VEVPGSQHIDSQ 82

RESULT 11
US-10-110-364-16
; Sequence 16, Application US/10110364
; GENERAL INFORMATION:
; APPLICANT: Handley, Harold H.
; APPLICANT: Haaparanta, Tapio
; APPLICANT: Ewalt, Karla L.
; TITLE OF INVENTION: ABS TOXIN B SUBUNIT MUTANTS WITH ALTERED
; TITLE OF INVENTION: CHEMICAL CONJUGATION CHARACTERISTICS
; FILE REFERENCE: ACTBIO.004A
; CURRENT APPLICATION NUMBER: US/10/110,364
; CURRENT FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: PCT/US00/27607
; PRIOR FILING DATE: 2000-10-05
; PRIOR APPLICATION NUMBER: 60/158,561
; PRIOR FILING DATE: 1999-10-08
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16
; LENGTH: 124
; TYPE: PRT
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(124)
; OTHER INFORMATION: LTB variant from NCBI gene bank GI: 1169505.
US-10-110-364-16

Query Match 100.0%; Score 62; DB 6; Length 124;
Best Local Similarity 100.0%; Pred. No. 0.00026;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VEVPGSQHIDSQ 12
|||||
Db 71 VEVPGSQHIDSQ 82

Db 71 VEVPGSQHIDSQ 82

RESULT 12
US-10-110-364-18
; Sequence 18, Application US/10110364
; GENERAL INFORMATION:
; APPLICANT: Handley, Harold H.
; APPLICANT: Haaparanta, Tapio
; APPLICANT: Ewalt, Karla L.
; TITLE OF INVENTION: ABS TOXIN B SUBUNIT MUTANTS WITH ALTERED
; TITLE OF INVENTION: CHEMICAL CONJUGATION CHARACTERISTICS
; FILE REFERENCE: ACTBIO.004A
; CURRENT APPLICATION NUMBER: US/10/110,364
; CURRENT FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: PCT/US00/27607
; PRIOR FILING DATE: 2000-10-05
; PRIOR APPLICATION NUMBER: 60/158,561
; PRIOR FILING DATE: 1999-10-08
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 18
; LENGTH: 124
; TYPE: PRT
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(124)
; OTHER INFORMATION: LTB variant from NCBI gene bank GI: 145833.
US-10-110-364-18

Query Match 100.0%; Score 62; DB 6; Length 124;
Best Local Similarity 100.0%; Pred. No. 0.00026;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VEVPGSQHIDSQ 12
|||||
Db 71 VEVPGSQHIDSQ 82

RESULT 13
US-10-110-364-19
; Sequence 19, Application US/10110364
; GENERAL INFORMATION:
; APPLICANT: Handley, Harold H.
; APPLICANT: Haaparanta, Tapio
; APPLICANT: Ewalt, Karla L.
; TITLE OF INVENTION: ABS TOXIN B SUBUNIT MUTANTS WITH ALTERED
; TITLE OF INVENTION: CHEMICAL CONJUGATION CHARACTERISTICS
; FILE REFERENCE: ACTBIO.004A
; CURRENT APPLICATION NUMBER: US/10/110,364
; CURRENT FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: PCT/US00/27607
; PRIOR FILING DATE: 2000-10-05
; PRIOR APPLICATION NUMBER: 60/158,561
; PRIOR FILING DATE: 1999-10-08
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 19
; LENGTH: 124
; TYPE: PRT
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(124)
; OTHER INFORMATION: LTB variant from NCBI gene bank GI: 1648865
US-10-110-364-19

Query Match 100.0%; Score 62; DB 6; Length 124;
Best Local Similarity 100.0%; Pred. No. 0.00026;

Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 VEPGSHIDSQ 12
 Db 71 VEPGSHIDSQ 82

RESULT 14

US-10-110-364-21
 ; Sequence 21, Application US/10110364
 ; GENERAL INFORMATION:
 ; APPLICANT: Handley, Harold H.
 ; APPLICANT: Haaparanta, Tapio
 ; APPLICANT: Ewalt, Karla L.
 ; TITLE OF INVENTION: AB5 TOXIN B SUBUNIT MUTANTS WITH ALTERED
 ; FILE REFERENCE: ACTBIO.004A
 ; CURRENT APPLICATION NUMBER: US/10/110,364
 ; CURRENT FILING DATE: 2002-04-05
 ; PRIOR APPLICATION NUMBER: PCT/US00/27607
 ; PRIOR FILING DATE: 2000-10-05
 ; PRIOR APPLICATION NUMBER: 60/158,561
 ; PRIOR FILING DATE: 1999-10-08
 ; NUMBER OF SEQ ID NOS: 31
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 21
 ; LENGTH: 124
 ; TYPE: PRT
 ; ORGANISM: Escherichia coli
 ; FEATURE:
 ; NAME/KEY: VARIANT
 ; LOCATION: (1)..(124)
 ; OTHER INFORMATION: LTB variant from NCBI gene bank GI: 408996.
 US-10-110-364-21

Query Match 100.0%; Score 62; DB 6; Length 124;
 Best Local Similarity 100.0%; Pred. No. 0.00026;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VEPGSHIDSQ 12
 Db 71 VEPGSHIDSQ 82

RESULT 15

US-10-110-364-23
 ; Sequence 23, Application US/10110364
 ; GENERAL INFORMATION:
 ; APPLICANT: Handley, Harold H.
 ; APPLICANT: Haaparanta, Tapio
 ; APPLICANT: Ewalt, Karla L.
 ; TITLE OF INVENTION: AB5 TOXIN B SUBUNIT MUTANTS WITH ALTERED
 ; FILE REFERENCE: ACTBIO.004A
 ; CURRENT APPLICATION NUMBER: US/10/110,364
 ; CURRENT FILING DATE: 2002-04-05
 ; PRIOR APPLICATION NUMBER: PCT/US00/27607
 ; PRIOR FILING DATE: 2000-10-05
 ; PRIOR APPLICATION NUMBER: 60/158,561
 ; PRIOR FILING DATE: 1999-10-08
 ; NUMBER OF SEQ ID NOS: 31
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 23
 ; LENGTH: 124
 ; TYPE: PRT
 ; ORGANISM: Escherichia coli
 ; FEATURE:
 ; NAME/KEY: VARIANT
 ; LOCATION: (1)..(124)
 ; OTHER INFORMATION: LTB variant from NCBI gene bank GI: 69630.
 US-10-110-364-23

Query Match 100.0%; Score 62; DB 6; Length 124;
 Best Local Similarity 100.0%; Pred. No. 0.00026;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 VEPGSHIDSQ 12
 Db 71 VEPGSHIDSQ 82

Search completed: July 3, 2002, 09:04:12
 Job time: 652 sec

Mon Jul 8 07:51:49 2002

us-09-786-648-3.rapn

Page 6

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 3, 2002, 08:56:32 ; Search time 46.57 Seconds

(without alignments)
24.760 Million cell updates/sec

Title: US-09-786-648-3

Sequence: 1 VEPYSGQHDSQ 12

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 2000000000Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

1: PIR.71.*
2: PIR1.*
3: PIR2.*
4: PIR3.*
5: PIR4.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysts of the total score distribution.

SUMMARIES

Result No.	Score	Match Length	DB ID	Description
1	62	100.0	124 1	XVVCB
2	62	100.0	124 1	GLECB
3	40	64.5	242 1	T34767
4	39	62.9	260 2	C96827
5	39	62.9	272 2	T14755
6	39	62.9	353 2	H75446
7	38	61.3	282 1	A35781
8	38	61.3	674 1	A41670
9	37	59.7	352 2	H85518
10	37	59.7	367 2	T01751
11	37	59.7	376 1	S17246
12	37	59.7	378 2	F90667
13	37	59.7	423 2	G96554
14	37	59.7	442 2	T39683
15	37	59.7	550 2	T01770
16	36	58.1	368 2	F98282
17	36	58.1	368 2	A83001
18	36	58.1	373 2	T47115
19	36	58.1	455 1	A69753
20	36	58.1	755 1	D95842
21	35	58.1	1785 2	S53976
22	35	56.5	142 2	S28757
23	35	56.5	259 2	B84125
24	35	56.5	284 2	F82031
25	35	56.5	302 2	JN0143
26	35	56.5	313 2	E88216
27	35	56.5	360 2	A90029
28	35	56.5	367 2	G83368
29	35	56.5	371 2	S23861

30	35	56.5	377 2	T29750
31	35	56.5	389 2	A86303
32	35	56.5	389 2	AE2980
33	35	56.5	427 2	JC5694
34	35	56.5	443 2	T29029
35	35	56.5	472 2	AG0432
36	35	56.5	539 2	S53529
37	35	56.5	593 2	T24379
38	35	56.5	670 2	T02092
39	35	56.5	960 2	A41638
40	35	56.5	961 2	A55380
41	35	56.5	1258 2	T29041
42	35	56.5	42	G71611
43	34	54.8	168 2	T27563
44	34	54.8	181 2	T45990
45	34	54.8	244 1	T01091

ALIGNMENTS

RESULT 1
XVVCB
cholera enterotoxin, B chain precursor VC1456 [validated] - Vibrio cholerae (strain N
M/Alternat names: enterotoxin beta chain
C/Species: Vibrio cholerae
C/Date: 24-Apr-1984 #sequence-revision 01-Sep-2000 #text-change 02-Feb-2001
C/Accession: S14624; S39238; S39241; H82196; JCI078; S17666; PC1010; A05130; A01819;
R/Dams, E.; de Wolf, M.; Dierick, W.
Submitted to the EMBL Data Library, March 1991
A/Description: Correction of the cholera toxin nucleotide sequence of the Vibrio chol
A/Reference number: S14624
A/Accession: S14624
A/Molecule type: DNA
A/Residues: 1-124 <DM>
A/Cross-references: EMBL:X58786; NID:q48420; PIDN:CAA41593.1; PID:q48422
A/Experimental source: Strain 2125
R/Lebens, M.; Holmgren, J.
Submitted to the EMBL Data Library, November 1993
A/Description: Structure and arrangement of the Cholera toxin genes in Vibrio Cholera
A/Reference number: S39238
A/Accession: S39238
A/Molecule type: DNA
A/Residues: 1-124 <LEB>
A/Cross-references: EMBL:X76390; NID:q433856; PIDN:CAA53973.1; PID:q433857
A/Accession: S39241
A/Molecule type: DNA
A/Residues: 1-124 <LEW>
A/Cross-references: EMBL:X76391; NID:q433859; PIDN:CAA53976.1; PID:q433861
R/Heidelberg, J.F.; Eelsen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.
Chadson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers
L. R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477/483, 2000
A/Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A/Reference number: A82035; MUID:20406833
A/Accession: H82196
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-124 <HEI>
A/Cross-references: GB:AE004224; GB:AE003852; NID:q9655952; PIDN:AAF4613.1; GSPDB:GN
A/Experimental source: serogroup O1, strain N16961; biotype El Tor
R/Shi, C.H.; Cao, C.; Zhang, J.S.; Ma, Q.J.
Chinese Biochem. J. 9, 395-399, 1993
A/Title: Nucleotide sequence analysis of the gene encoding the classical biotype V.ch
A/Reference number: JCI078
A/Accession: JCI078
A/Molecule type: DNA
A/Residues: 1-20, 'Q', 22-31, 'Q', 33-38, 'H', 40-49, 'G', 51-67, 'T', 69-124 <SHI>
A/Experimental source: classical biotype strain 569b
R/Dams, E.; de Wolf, M.; Dierick, W.
Biochim. Biophys. Acta 1090, 139-141, 1991
A/Title: Nucleotide sequence analysis of the CT operon of the Vibrio cholerae class
A/Reference number: S17665; MUID:91355224

A:Accession: S17666
 A:Molecule type: DNA
 A:Residues: 1-38,'H','40-67','T','69-124 <DA2>
 A:Cross-references: EMBL:X58785; NID:948886; PIDN:CAA41591.1; PID:948890
 R:Ma, Q.J.; Liu, C.X.; Xiong, L.S.; Yu, X.Q.
 Sci. Sin. B Chem. Biol. Agric. Med. Earth Sci. 34, 274-280, 1991
 A:Title: B subunit of cholera toxin produced in *Escherichia coli*.
 A:Reference number: PC1010
 A:Accession: PC1010
 A:Molecule type: protein
 A:Residues: 22-38,'H','40-41 <MAO>
 R:Mekeanos, J.J.; Swartz, D.J.; Pearson, G.D.N.; Hartford, N.; Groyne, F.; de Wilde, M.
 Nature 306, 551-557, 1983
 A:Reference number: A93320; MWID:84068199
 A:Accession: A05130
 A:Molecule type: DNA
 A:Residues: 1-32,'S','34-74','S','76-124 <MEK>
 A:Cross-references: GB:X00171; NID:948347; PIDN:CAA2496.1; PID:g758351
 R:Kurosky, A.; Markel, D.E.; Peterson, J.W.
 J. Biol. Chem. 252, 7257-7264, 1977
 A:Title: Covalent structure of the beta chain of cholera enterotoxin.
 A:Reference number: A01819; MWID:78005537
 A:Accession: A01819
 A:Molecule type: protein
 A:Residues: 22-38,'H','40-42','N','44-67','T','69-90','N','92-124 <KUR>
 R:Uai, C.Y.
 J. Biol. Chem. 252, 7249-7256, 1977
 A:Title: Determination of the primary structure of cholera toxin B subunit.
 A:Reference number: A38033; MWID:78005536
 A:Accession: A38033
 A:Molecule type: protein
 A:Residues: 22-38,'H','40-42','N','44-67','T','69','E','71-90','N','92-124 <LA1>
 A:Note: the difference at residue 70 may be due to deamidation during preparation
 R:Nakshima, Y.; Napietkowski, P.; Schaffer, D.E.; Konigsberg, W.H.
 FEBS Lett. 68, 275-278, 1976
 A:Title: Primary structure of the B subunit of cholera enterotoxin.
 A:Reference number: A38034; MWID:77026365
 A:Accession: A38034
 A:Molecule type: protein
 A:Residues: 22-38,'H','40-67','T','69','E','71','Q','74-75','V','78-86','Q','88-99','Q','101-103','Q'
 R:Takao, T.; Watanabe, H.; Shimonishi, Y.
 Eur. J. Biochem. 146, 503-508, 1985
 A:Title: Facile identification of protein sequences by mass spectrometry.
 A:Reference number: A21910; MWID:85126976
 A:Accession: A21910
 A:Molecule type: protein
 A:Residues: 22-38,'H','40-42','N','44-67','T','69-90','N','92-124 <TAK>
 A:Experimental source: biotype Inaba 569B
 A:Note: Asn-65 was partially deaminated to Asp
 C:Comment: The authors translated the codon TCA for residue 33 as Tyr.
 C:Genetics:
 A:Gene: VC1456
 A:Map position: 1
 C:Complex: the cholera enterotoxin molecule contains three kinds of chains; an alpha and
 ciate noncovalently with the subunit B, an aggregate of five beta chains
 C:Function:
 A:Description: involved in binding of the toxin to cell membranes
 C:Superfamily: cholera enterotoxin beta chain
 C:Keywords: enterotoxin; toxin
 F:1-21/Domain: signal sequence #status predicted <SIG>
 F:22-124/Product: cholera enterotoxin chain B #status predicted <MAT>
 F:30-107/Disulfide bonds: #status experimental

Query Match 100.0%; Score 62; DB 1; Length 124;
 Best Local Similarity 100.0%; Pred. No. 0.00019;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VEVPGSOHDSQ 12
 |||||||||
 Db 71 VEVPGSOHDSQ 82

RESULT 2
 QLECB
 heat-labile enterotoxin chain B precursor - *Escherichia coli*
 C:Species: *Escherichia coli*
 C:Date: 29-Jun-1981 #sequence, revision 29-Jun-1981 #text, change 18-Jun-1999
 C:Accession: A01820; B26946; I41194; I41287; I67644; A61475
 R:Dallas, W.S.; Falkow, S.
 Nature 288, 499-501, 1980
 A:Title: Amino acid sequence homology between cholera toxin and *Escherichia coli* heat
 A:Reference number: A01820; MWID:81074965
 A:Accession: A01820
 A:Molecule type: mRNA
 A:Residues: 1-124 <DAL>
 R:Yamamoto, T.; Gojobori, T.; Yokota, T.
 J. Bacteriol. 169, 1352-1357, 1987
 A:Title: Evolutionary origin of pathogenic determinants in enterotoxigenic *Escherichia*
 A:Reference number: A26946; MWID:87137303
 A:Accession: B26946
 A:Molecule type: DNA
 A:Residues: 1-27,'E','29-63','K','65-124 <YAM>
 A:Cross-references: EMBL:M15363; NID:9148335; PIDN:AAA24792.1; PID:g148336
 R:Leong, J.; Vinal, A.C.; Dallas, W.S.
 Infect. Immun. 48, 73-77, 1985
 A:Title: Nucleotide sequence comparison between heat-labile toxin B-subunit cistrons
 A:Reference number: I41194; MWID:85156481
 A:Accession: I41194
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-5,'F','7-17','C','19-24','S','26-27','E','29-33','H','35-63','K','65-66','A','68-122'
 A:Cross-references: GB:M17874; NID:g145830; PIDN:AAA8064.1; PID:g145831
 A:Experimental source: plasmid ENT-R PCG86
 R:Ibrahim, I.; Gentz, R.
 J. Biol. Chem. 262, 10189-10194, 1987
 A:Title: A functional interaction between the signal peptide and the translation appa
 ticulum.
 A:Reference number: I41287; MWID:87280041
 A:Accession: I41287
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-22 <RE2>
 A:Cross-references: GB:M17101; NID:g146375; PIDN:AAA23973.1; PID:g146376
 R:Inoue, T.; Tsuji, T.; Koto, M.; Imamura, S.; Miyama, A.
 FEBS Microbiol. Lett. 108, 157-161, 1993
 A:Title: Amino acid sequence of heat-labile enterotoxin from chicken enterotoxigenic
 A:Reference number: I53542; MWID:93252225
 A:Accession: I57644
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-17,'C','19','Y','21-24','S','26-27','E','29-63','K','65-66','A','68-122','E','124 <R
 A:Cross-references: GB:S60731; NID:g408994; PIDN:AAC6044.1; PID:g408996
 R:Tsuji, T.; Lida, T.; Honda, T.; Mawatani, T.; Negahama, M.; Sakurai, J.; Wada, K.;
 Microb. Pathog. 2, 381-390, 1987
 A:Title: A unique amino acid sequence of the B subunit of a heat-labile enterotoxin 1
 A:Reference number: A61475; MWID:89180953
 A:Accession: A61475
 A:Molecule type: protein
 A:Residues: 22-24,'S','26-27','E','29-63','K','65-66','A','68-95','A','97-122','E','124 <TSU>
 A:Experimental source: strain 240-3
 C:Complex: the heat-labile enterotoxin molecule contains one A chain and five or six
 C:Function:
 A:Description: the biological activity of the toxin is produced by the A chain, which
 C:Superfamily: cholera enterotoxin beta chain
 C:Keywords: enterotoxin
 F:1-21/Domain: signal sequence #status predicted <SIG>
 F:22-124/Product: heat-labile enterotoxin chain B #status predicted <MAT>
 F:30-107/Disulfide bonds: #status predicted

Query Match 100.0%; Score 62; DB 1; Length 124;
 Best Local Similarity 100.0%; Pred. No. 0.00019;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VEVPGSOHDSQ 12

Db 71 VEVPGSOHIDSQ 82

RESULT 3

T34767
hypothetical protein SC2A11.21c SC2A11.21c - Streptomyces coelicolor
C:Species: Streptomyces coelicolor
C:Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 05-Nov-1999
C:Accession: T34767
R:Murphy, L.; Harris, D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, August 1998
A:Reference number: 221556
A:Accession: T34767
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-242 <MUR>
A:Cross-references: EMBL:AL031184; PIDN:CAA20190.1; GSPDB:GN00070; SCOEDB:SC2A11.21c
C:Genetics:
A:gene: SCOEDB:SC2A11.21c

Query Match 64.5%; Score 40; DB 2; Length 242;
Best Local Similarity 70.0%; Pred. No. 6.1;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 VEVPGSOHID 10
|||||:|
Db 204 VEVPGTDHFD 213

RESULT 4

C96827
protein P20B17.2 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C:Accession: C96827
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
anssen, N.F.; Hughes, B.; Hultzer, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marzalli,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719
A:Accession: C96827
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-260 <STO>
A:Cross-references: GB:AE005173; NID:g7715588; PIDN:AAF6106.1; GSPDB:GN00141
C:Genetics:
A:gene: P20B17.2
A:Map position: 1

Query Match 62.9%; Score 39; DB 2; Length 260;
Best Local Similarity 50.0%; Pred. No. 10;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

OY 1 VEVPGSOHIDSQ 12
|||||:|
Db 65 VEVPGKOHVSEK 76

RESULT 5

T14755
hypothetical protein DKFZp564A0122.1 - human
C:Species: Homo sapiens (man)
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999

C:Accession: T14755
R:Mambuli, R.; Heubner, D.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, August 1999
A:Reference number: Z18181
A:Accession: T14755
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-272 <MAN>
A:Cross-references: EMBL:AL110209
A:Experimental source: fetal brain; clone DKFZp564A0122
C:Genetics:
A:Note: DKFZp564A0122.1

Query Match 62.9%; Score 39; DB 2; Length 272;
Best Local Similarity 66.7%; Pred. No. 11;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 2 VEVPGSOHID 10
|||||:|
Db 246 ELPGESEHIE 254

RESULT 6

H75446
(S)-2-hydroxy-acid oxidase - Deinococcus radiodurans (strain R1)
C:Species: Deinococcus radiodurans
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 28-Jul-2000
C:Accession: H75446
R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.
M.; Shen, M.; Yamathavan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.;
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A:Reference number: A75250; MUID:20036896
A:Accession: H75446
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-353 <MHI>
A:Cross-references: GB:AE001954; GB:AE000513; NID:g6458751; PIDN:AAF10604.1; PID:g645
A:Experimental source: strain R1
C:Genetics:
A:gene: DR1031
A:Map position: 1
C:Superfamily: (S)-2-hydroxy-acid oxidase; (S)-2-hydroxy-acid oxidase homology
F;3-297/Domain: (S)-2-hydroxy-acid oxidase homology <2HY>

Query Match 62.9%; Score 39; DB 2; Length 353;
Best Local Similarity 75.0%; Pred. No. 14;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 3 VPGSOHID 10
|||||:|
Db 189 VPGSEHID 196

RESULT 7

A35781
hippocampus-derived neurotrophic factor precursor - rat
N:Altenate names: neurotrophin-3 precursor
C:Species: Rattus norvegicus (Norway Rat)
C:Date: 05-Oct-1990 #sequence_revision 05-Oct-1990 #text_change 16-Jul-1999
C:Accession: A35781; A40094
R:Ernfors, P.; Ibanez, C.F.; Ebendal, T.; Olson, L.; Persson, H.
Proc. Natl. Acad. Sci. U.S.A. 87, 5454-5458, 1990
A:Title: Molecular cloning and neurotrophic activities of a protein with structural s
A:Reference number: A35781; MUID:90319130
A:Accession: A35781
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-282 <ERN>
A:Cross-references: GB:M34643

R;Maisonnier, P.C.; Belluscio, L.; Squinto, S.; Ip, N.Y.; Firth, M.E.; Lindsay, R.M.; Science 247, 1446-1451, 1990
A:Title: Neurotrophin-3: a neurotrophic factor related to NGF and BDNF.
A:Reference number: A40094; MUID:90208301
A:Accession: A40094
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 25-282 <MAI>
A:Cross-references: GB:M33968; NID:g205771; PIDN:AAA1727.1; PID:g205772
C:Superfamily: nerve growth factor beta chain

Query Match 61.3%; Score 38; DB 2; Length 282;
Best Local Similarity 54.5%; Pred. No. 17;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
OY 1 VEPGSHIDS 11
Db 1 VDPGNSHTDA 11

RESULT 8
A41670
Carbon-monoxide dehydrogenase (EC 1.2.99.2) beta chain [similarity] - Clostridium thermac
C:Species: Clostridium thermacellum
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 28-Jul-2000
C:Accession: A41670
R;Morton, T.A.; Runquist, J.A.; Ragsdale, S.W.; Shanmugasundaram, T.; Wood, H.G.; Ljung
J. Biol. Chem. 266, 23824-23828, 1991
A:Title: The primary structure of the subunits of carbon monoxide dehydrogenase/acetyl-C
A:Reference number: A41670; MUID:92084676
A:Accession: A41670
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-674 <MOR>
A:Cross-references: GB:M62727; NID:g144784; PIDN:AAA23228.1; PID:g144785
C:Superfamily: carbon-monoxide dehydrogenase beta chain; hybrid cluster [4Fe-2S-30] homoc
C:Keywords: 4Fe-2S-30; 4Fe-4S; electron transfer; iron-sulfur protein; metalloprotein; c
F:256-606/Domain: hybrid cluster [4Fe-2S-30] homology <HCL>
F:68,71,76,90/Binding site: 4Fe-4S cluster (Cys) (covalent) #status predicted
F:283,321,355,470,500,550,585/Binding site: Ni-3Fe-2S-30 cluster (His, Glu, Cys, Cys, Cy
F:470/Modified site: cysteine persulfide (Cys) #status predicted

Query Match 61.3%; Score 38; DB 1; Length 674;
Best Local Similarity 54.5%; Pred. No. 45;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 2 EVPGSHIDSQ 12
Db 379 KPGAYHIDYQ 389

RESULT 9
B85518
hypothetical protein Z0347 [imported] - Escherichia coli (strain O157:H7, substrain EDL8
C:Species: Escherichia coli
C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
C:Accession: B85518
R;Penna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
Hiller, L.; Grobbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamoudis, K.; Apodaca,
Nature 409, 529-533, 2001
A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A:Reference number: A85480; MUID:21074935; PMID:11206551
A:Accession: B85518
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-352 <STO>
A:Cross-references: GB:AE005174; NID:g12513064; PIDN:AAG54606.1; GSPDB:GN00145; UMGP:Z03
C:Genetics:
A:Experimental source: strain O157:H7, substrain EDL933
A:Gene: Z0347

Query Match 59.7%; Score 37; DB 2; Length 352;
Best Local Similarity 60.0%; Pred. No. 35;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 VEPGSHID 10
Db 320 VVPGANHVD 329

RESULT 10
T01751
gibberellin 20-oxidase - common tobacco
N:Alternate names: Ntcl6 protein
C:Species: Nicotiana tabacum (common tobacco)
C:Date: 19-Feb-1999 #sequence_revision 19-Feb-1999 #text_change 20-Jun-2000
C:Accession: T01751
R;Tanaka-Ueguchi, M.; Itoh, H.; Oyama, N.; Koshioka, M.; Matsuo, M.
submitted to the EMBL Data Library, July 1998
A:Description: Over-expression of a tobacco homeobox gene, NTH15, decreases the expe
A:Reference number: Z14418
A:Accession: T01751
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-367 <TAN>
A:Cross-references: EMBL:AB016084
C:Genetics:
A:Gene: Ntcl6
C:Superfamily: 1-aminocyclopropane-1-carboxylate oxidase

Query Match 59.7%; Score 37; DB 2; Length 367;
Best Local Similarity 87.5%; Pred. No. 36;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 EVPGSHI 9
Db 149 EVPGSHI 156

RESULT 11
S17246
choistmate synthase (EC 4.6.1.4) - yeast (Saccharomyces cerevisiae)
N:Alternate names: protein G2501; protein YGL148W
C:Species: Saccharomyces cerevisiae
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 16-Jun-2000
C:Accession: S17246; S64162
R;Jones, D.G.L.; Reusser, U.; Braus, G.H.
Mol. Microbiol. 5, 2143-2152, 1991
A:Title: Molecular cloning, characterization and analysis of the regulation of the AR
A:Reference number: S17246; MUID:92114793
A:Accession: S17246
A:Molecule type: DNA
A:Residues: 1-376 <JON>
A:Cross-references: EMBL:X60190; NID:g3386; PIDN:CAA42745.1; PID:g3387
R;Volckert, G.; Voel, M.; Vernasselt, P.; Defoor, E.
submitted to the Protein Sequence Database, May 1996
A:Reference number: S64153
A:Accession: S64162
A:Molecule type: DNA
A:Residues: 1-376 <VOL>
A:Cross-references: EMBL:Z72670; NID:g1322731; PIDN:CAA96860.1; PID:g1322732; GSPDB:G
A:Experimental source: strain S288C
C:Genetics:
A:Gene: SGD:ARO2; MIPS:YGL148W
A:Cross-references: SGD:S0003116; MIPS:YGL148W
A:Map position: 7L
C:Superfamily: choistmate synthase
C:Keywords: phosphorus-oxygen lyase; transmembrane protein
F:347-363/Domain: transmembrane #status predicted <TMW>

Query Match 59.7%; Score 37; DB 1; Length 376;

Best Local Similarity 70.0%; Pred. No. 37;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 VEPGSOHID 10
Db 264 VVPGANHVD 273

RESULT 12

hypothetical protein ECS0310 [imported] - Escherichia coli (strain O157:H7, substrain R1)
C:Species: Escherichia coli
C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 18-Jul-2001
C:Accession: F90667
R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.; Gatawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genomic reference number: A99629; MID:21156231; PMID:11258796
A:Accession: F90667
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-378 <HAY>
A:Cross-references: GB:BA000007; PIDN:BA033733.1; PID:q13359767; GSPDB:GN00154
A:Experimental source: strain O157:H7, substrain R1MD 0509952
C:Genetics:
A:Gene: ECS0310

Query Match 59.7%; Score 37; DB 2; Length 378;
Best Local Similarity 60.0%; Pred. No. 37;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 VEPGSOHID 10
Db 346 VVPGANHVD 355

RESULT 13

G96554
hypothetical protein F19C24.16 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C:Accession: G96554
R:Theologos, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; Jensen, N.F.; Hughes, B.; Hultzer, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, Ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MID:21016719
A:Accession: G96554
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-423 <STO>
A:Cross-references: GB:AE005173; MID:g11094753; PIDN:AGC29686.1; GSPDB:GN00141
C:Genetics:
A:Gene: F19C24.16
A:Map position: 1

Query Match 59.7%; Score 37; DB 2; Length 423;
Best Local Similarity 75.0%; Pred. No. 42;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 3 VPGSOHID 10
Db 403 VPGIOHVD 410

RESULT 14

T39683
zucotin-like protein - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C:Date: 20-Oct-2000 #sequence_revision 20-Oct-2000 #text_change 20-Oct-2000
C:Accession: T39683; T40195
R:Oliver, K.; Harris, D.; Wood, V.; Rajandream, M.A.; Barrell, B.G.
Submitted to the EMBL Data Library, March 1998
A:Reference number: Z21869
A:Accession: T39683
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-124 <OLI>
A:Cross-references: EMBL:AL049489; PIDN:CAB39796.1; GSPDB:GN00067; SPDB:SPBC1778.01c
A:Experimental source: strain 972h-; cosmid c1778
R:Wood, V.; Rajandream, M.A.; Barrell, B.G.; Lauber, J.; Hilbert, H.; Duesterhoeft, A.
Submitted to the EMBL Data Library, February 1998
A:Reference number: Z21910
A:Accession: T40195
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 89-442 <WOO>
A:Cross-references: EMBL:Z97992; PIDN:CAB10796.1; GSPDB:GN00067; SPDB:SPBC30D10.01
C:Genetics:
A:Gene: SPDB:SPBC1778.01c; SPDB:SPBC30D10.01
A:Map position: 2

Query Match 59.7%; Score 37; DB 2; Length 442;
Best Local Similarity 50.0%; Pred. No. 44;
Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

OY 2 EVPGSOHIDS 11
Db 371 DVPSAHERVDS 380

RESULT 15

T01770
hypothetical protein A_IG002P16.9 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 19-Feb-1999 #sequence_revision 19-Feb-1999 #text_change 22-Oct-1999
C:Accession: T01770
R:Miller, N.; Beck, C.; Kramer, J.
Submitted to the EMBL Data Library, June 1997
A:Description: The sequence of A. thaliana IG002P16.
A:Reference number: Z14421
A:Accession: T01770
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-550 <ML>
A:Cross-references: EMBL:AF007270; MID:g2191157; PID:g2191172; GSPDB:GN00063; ATSP:A_IG002P16.9
A:Gene: ATSP:A_IG002P16.9
A:Map position: 5
A:Introns: 159/1; 272/1; 434/2; 477/3

Query Match 59.7%; Score 37; DB 2; Length 550;
Best Local Similarity 66.7%; Pred. No. 56;
Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 1 VEPGSOHIDSQ 12
Db 260 VTCSGSOHIDFQ 271

Search completed: July 3, 2002, 08:56:32
Job time: 232 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 3, 2002, 09:04:41 ; Search time 21.51 Seconds

(without alignments)
21.601 Million cell updates/sec

Title: US-09-786-648-3
Perfect score: 62
Sequence: 1 VEPGSGHDSQ 12

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues
Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	62	100.0	124	1 CHTB_VIBCH	P01556 vibrio chol
2	62	100.0	124	1 ELBH_ECOLI	P13811 escherichia
3	62	100.0	124	1 EUBP_ECOLI	P32890 escherichia
4	38	61.3	674	1 DCMB_MOOTH	P27989 moorella th
5	37	59.7	376	1 AROC_YEAST	P28777 saccharomyc
6	36	58.1	240	1 YB5A_THEMEA	P58009 thermotoga
7	36	58.1	455	1 GUDH_BACSU	P42238 bacillus su
8	36	58.1	743	1 BGAL_THERT	P77989 thermoaer
9	36	58.1	855	1 SRI4_HUMAN	O94596 homo sapien
10	36	58.1	1785	1 GLS3_YEAST	O04952 saccharomyc
11	35	56.5	302	1 PHEB_PSESP	P31019 pseudomonas
12	35	56.5	313	1 YP68_CAEBL	O09217 caenorhabdi
13	35	56.5	371	1 YM18_PSEAE	O01609 pseudomonas
14	35	56.5	539	1 TYRO_ASPOR	O00234 aspergillus
15	35	56.5	670	1 INVI_MAIZE	P49175 zea mays (m
16	35	56.5	960	1 CHS3_NEUCR	P28070 neurospora
17	35	56.5	960	1 FGD1_MOUSE	P52734 mus musculu
18	35	56.5	961	1 FGD1_HUMAN	P98174 homo sapien
19	34	54.8	219	1 CIDA_HUMAN	O60543 homo sapien
20	34	54.8	244	1 NUHM_ARATH	O22769 arabidopsis
21	34	54.8	365	1 FTZ3_PYRAB	O94075 pyrococcus
22	34	54.8	365	1 FTZ3_PYRHO	O50060 pyrococcus
23	34	54.8	419	1 GSC_DROME	O13666 drosophila
24	34	54.8	432	1 AROC_NEUCR	O12640 neurospora
25	34	54.8	456	1 SR54_THEAC	O9hkt0 thermoplas
26	34	54.8	459	1 IL7R_MOUSE	P16872 mus musculu
27	34	54.8	491	1 CD5_RAT	P51882 rattus norv
28	34	54.8	500	1 GABT_HUMAN	P80404 homo sapien
29	34	54.8	508	1 Y202_HUMAN	O92599 homo sapien
30	34	54.8	557	1 HLYB_SERMA	P15321 serattia ma
31	34	54.8	560	1 INRI_SHEEP	O28589 ovvis aries
32	34	54.8	775	1 THIL_SCHPO	P36598 schizosacch
33	34	54.8	842	1 DPOL_HPBVM	P31870 hepatitis b

ALIGNMENTS

RESULT 1	ID	CHTB_VIBCH	STANDARD:	PRT:	124 AA.
AC	P01556	Q9J002			
DT	21-JUL-1986	(Rel. 01, Created)			
DT	13-AUG-1987	(Rel. 05, Last sequence update)			
DT	16-OCT-2001	(Rel. 40, Last annotation update)			
DE	Cholera enterotoxin, beta chain precursor.				
GN	CTXB OR TOXB OR VC1456.				
OS	Vibrio cholerae.				
OC	Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.				
OX	NCBI_TaxID=666;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=84061784; PubMed=6315707;				
RA	Lockman H., Kaper J.B.;				
RT	"Nucleotide sequence analysis of the A2 and B subunits of Vibrio				
RT	cholerae enterotoxin.";				
RL	J. Biol. Chem. 258:13722-13726(1983).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RX	STRAIN=EL TOR 2125;				
RC	MEDLINE=84068199; PubMed=6646234;				
RA	Mekalanos J.J., Swartz D.J., Pearson G.D.N., Harford N., Groyne F.,				
RA	de Wilde M.;				
RT	"Cholera toxin genes: nucleotide sequence, deletion analysis and				
RT	vaccine development.";				
RL	Nature 306:551-557(1983).				
RN	[3]				
RP	SEQUENCE FROM N.A.				
RA	STRAIN=EL TOR 2125;				
RC	Dans E., de Wolf M., Dierick W.;				
RA	Submitted (MAY-1991) to the EMBL/GenBank/DBJ databases.				
RN	[4]				
RP	SEQUENCE FROM N.A.				
RX	STRAIN=4260B / SEROTYPE O139;				
RC	MEDLINE=94237453; PubMed=8181723;				
RA	Lebens M., Holmgren J.;				
RT	"Structure and arrangement of the cholera toxin genes in Vibrio				
RT	cholerae O139.";				
RL	FEMS Microbiol. Lett. 117:197-202(1994).				
RN	[5]				
RP	SEQUENCE FROM N.A.				
RA	STRAIN=1854 / O139-BENGAL;				
RC	Yanamoto K., Do V.G.R.F., Xu M., Iida T., Miwatani T., Albert M.J.,				
RA	Honda T.;				
RL	Submitted (MAY-1994) to the EMBL/GenBank/DBJ databases.				
RN	[6]				
RP	SEQUENCE FROM N.A.				
RA	STRAIN=EL TOR N16961 / SEROTYPE O1;				
RC	MEDLINE=20406833; PubMed=10952301;				
RA	Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwin M.L.,				
RA	Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,				
RA	Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,				
RA	Ernolova M.D., Vamathevan J., Bass S., Qin H., Dragoli I., Sellers P.,				
RA	McDonald L., Uitterback T., Fleischmann R.D., Niernan W.C., White O.,				

34	34	54.8	1151	1	Y245_TREPA	O83273 treponema p
35	33.5	54.0	375	1	MASP_MOUSE	P70124 mus musculu
36	33.5	54.0	375	1	MASP_RAT	P70564 rattus norv
37	33	53.2	259	1	MOB2_YEAST	P43563 saccharomyc
38	33	53.2	264	1	FTSQ_YEAST	O86038 streptomyc
39	33	53.2	291	1	CORC_BUCAI	P57518 buchiera ap
40	33	53.2	386	1	YVAN_BACSU	P37535 bacillus su
41	33	53.2	454	1	VNUC_THOCH	P89216 thogoto vir
42	33	53.2	504	1	A37C_DROME	P18487 drosophila
43	33	53.2	534	1	YOG1_CAEBL	P34610 caenorhabdi
44	33	53.2	621	1	HEM1_AGABI	O92403 agaricus bi
45	33	53.2	666	1	PD14_MOUSE	O92183 mus musculu

RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
 RA Fraser C.M.;
 RT "DNA sequence of both chromosomes of the cholera pathogen *Vibrio*
 RT *cholerae*.";
 RL Nature 406:477-483(2000).
 RN [7]
 RP SEQUENCE OF 22-124.
 RX MEDLINE=78005357; PubMed=903363;
 RA Kurosky A., Markel D.E., Peterson J.W.;
 RT "Covalent structure of the beta chain of cholera enterotoxin.";
 RL J. Biol. Chem. 252:7257-7264(1977).
 RN [8]
 RP SEQUENCE OF 22-124.
 RX MEDLINE=78005356; PubMed=903362;
 RA Lai C.-Y.;
 RT "Determination of the primary structure of cholera toxin B subunit.";
 RL J. Biol. Chem. 252:7249-7256(1977).
 RN [9]
 RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS).
 RX MEDLINE=9427319; PubMed=8003954;
 RA Merritt E.A., Sarfaty S., van den Akker F., L'Hoir C., Martiat J.A.,
 RA Hol W.G.J.;
 RT "Crystal structure of cholera toxin B-pentamer bound to receptor GMI
 RT pentasaccharide.";
 RL Protein Sci. 3:166-175(1994).
 RN [10]
 RP X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS).
 RX MEDLINE=95387394; PubMed=7658472;
 RA Zhang R.G., Westbrook M.L., Westbrook E.M., Scott D.L., Orlowski Z.,
 RA Maulik P.R., Reed R.A., Shipley G.G.;
 RT "The 2.4 A crystal structure of cholera toxin B subunit pentamer:
 RT choleragenoid.";
 RL J. Mol. Biol. 251:550-562(1995).
 RN [11]
 RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).
 RC STRAIN-OGAWA 41 / CLASSICAL BIOTYPE;
 RX MEDLINE=97376625; PubMed=9232653;
 RA Merritt E.A., Sarfaty S., Jobling M.G., Chang T., Holmes R.K.,
 RA Hirst T.R., Hol W.G.J.;
 RT "Structural studies of receptor binding by cholera toxin mutants.";
 RL Protein Sci. 6:1516-1528(1997).
 CC -1- FUNCTION: THE BETA CHAIN AGGREGATE (B SUBUNIT) IS INVOLVED IN
 CC BINDING TO CELL MEMBRANES.
 CC -1- SUBUNIT: CONTAINS 3 KINDS OF CHAINS. AN ALPHA AND A GAMMA CHAIN
 CC (FROM THE SAME PRECURSOR MOLECULE), LINKED BY AN INTERCHAIN
 CC DISULFIDE BOND, ASSOCIATE NONCOVALENTLY WITH AN AGGREGATE OF 4 TO
 CC 6 BETA CHAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X00171; CAA24996.1; -
 DR EMBL; K01170; AAA27573.1; -
 DR EMBL; D30053; BAA06291.1; -
 DR EMBL; X58786; CAA1593.1; -
 DR EMBL; X76380; CAA53973.1; -
 DR EMBL; X76391; CAA53976.1; -
 DR EMBL; AE04224; AAF94613.1; -
 DR PIR; A01819; XVCB.
 DR PIR; A05130; A05130.
 DR PIR; S14624; S14624.
 DR PDB; 2CHB; 03-DEC-97.
 DR PDB; 3CHB; 12-AUG-98.
 DR PDB; 1CHP; 08-MAR-96.
 DR PDB; 1CHO; 08-MAR-96.
 DR PDB; 1FGB; 23-DEC-96.
 DR PDB; 1XPB; 01-APR-97.
 DR PDB; 1XTC; 01-AUG-96.

DR PDB; 1CTL; 15-OCT-97.
 DR TIGR; VC1456; -
 DR InterPro: IPR001835; Enterotoxin_B.
 DR Pfam: PF01376; Enterotoxin_B.1.
 DR PRINTS: PR00772; ENTEROTOXINB.
 DR ProDom: PD012805; Enterotoxin_B.1.
 KW Membrane; Enterotoxin; Signal; 3D-structure; Complete proteome.
 FT SIGNAL 1 21
 FT CHAIN 22 124 CHOLERA ENTEROTOXIN, BETA CHAIN.
 FT DISULFID 30 107
 FT CONFLICT 33 33
 FT CONFLICT 39 39 Y -> S (IN REF. 2).
 FT CONFLICT 43 43 Y -> H (IN REF. 7 AND 8).
 FT CONFLICT 68 68 I -> T (IN REF. 7 AND 8).
 FT CONFLICT 70 70 Q -> E (IN REF. 8).
 FT CONFLICT 75 75 G -> S (IN REF. 2).
 FT CONFLICT 91 91 D -> N (IN REF. 7 AND 8).
 FT HELIX 26 30
 FT TURN 31 32
 FT TURN 34 35
 FT STRAND 36 44
 FT STRAND 47 51
 FT TURN 54 55
 FT STRAND 58 62
 FT TURN 64 65
 FT STRAND 68 71
 FT TURN 76 77
 FT HELIX 80 99
 FT TURN 100 100
 FT STRAND 102 109
 FT STRAND 115 123
 SQ SEQUENCE 124 AA; 13957 MW; 9AA393E3EA8E3EBF CRC64;

Query Match 100.0%; Score 62; DB 1; Length 124;
 Best Local Similarity 100.0%; Pred. No. 6e-05;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 VEVPGSQHDSQ 12
 DB 71 VEVPGSQHDSQ 82

RESULT 2
 ID EMBL; ECOLI STANDARD; PRT; 124 AA.
 AC P13811;
 DT 01-JAN-1990 (Rel. 13, Created)
 DT 01-JAN-1990 (Rel. 13, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Heat-labile enterotoxin B chain precursor (LT-B, human) (LTH-B).
 GN EYB OR LYB.
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Escherichia.
 OC NCBI_TaxID=562;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ISOLATE H74-114;
 RX MEDLINE=85156481; PubMed=3884513;
 RA Leong J., Vinal A.C., Dallas W.S.;
 RT "Nucleotide sequence comparison between heat-labile toxin B-subunit
 RT cistrons from Escherichia coli of human and porcine origin.";
 RL Infect. Immun. 48:73-77(1985).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ISOLATE H10407;
 RX MEDLINE=83114628; PubMed=6759877;
 RA Yamamoto T., Tamura T.A., Yokota T., Takano T.;
 RT "Overlapping genes in the heat-labile enterotoxin operon originating
 RT from Escherichia coli human strain.";
 RL Mol. Gen. Genet. 188:356-359(1982).
 RN [3]

RP SEQUENCE FROM N.A.
 RC STRAIN-ISOLATE H10407;
 RX MEDLINE=93252225; PubMed=8486242;
 RA Inoue T., Tsuji T., Koto M., Imamura S., Miyama A.;
 RT "Amino acid sequence of heat-labile enterotoxin from chicken
 RT enterotoxigenic Escherichia coli is identical to that of human strain
 RT H 10407.";
 RL FEBS Microbiol. Lett. 108:157-161(1993).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ETEC LT 87;
 RA Germani Y., Desperrier J.M.;
 RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP DISCUSSION OF SEQUENCE.
 RX MEDLINE=95349400; PubMed=7623669;
 RA Domenighini M., Plaza M., Jobling M.G., Holmes R.K., Rappuoli R.;
 RT "Identification of errors among database sequence entries and
 RT comparison of correct amino acid sequences for the heat-labile
 RT enterotoxins of Escherichia coli and Vibrio cholerae.";
 RL Mol. Microbiol. 13:1165-1167(1995).
 RN [6]
 RP X-RAY CRYSTALLOGRAPHY (3.04 ANGSTROMS).
 RX MEDLINE=99185101; PubMed=10085117;
 RA Metkovic-Galogovic D., Iloregian A., D'Aacunto M.R., Battistutta R.,
 RA Tossi A., Palu G., Zanotti G.;
 RT "Crystal structure of the B subunit of Escherichia coli heat-labile
 RT enterotoxin carrying peptides with anti-herpes simplex virus type 1
 RT activity.";
 RL J. Biol. Chem. 274:8764-8769(1999).
 CC -1- FUNCTION: THE BIOLOGICAL ACTIVITY OF THE TOXIN IS PRODUCED BY
 CC THE A CHAIN, WHICH ACTIVATES INTRACELLULAR ADENYL CYCLASE.
 CC -1- SUBUNIT: HETEROHEXAMER OF ONE A CHAIN AND OF FIVE B CHAINS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC -----
 DR EMBL: M17874; AAA98064.1; -;
 DR EMBL: J01646; AAB02982.1; -;
 DR EMBL: S60731; AAC60441.1; -;
 DR EMBL: X83966; CAA58800.1; -;
 DR PDB: 1LTR; 23-MAR-99.
 DR InterPro: IPR001835; Enterotoxin_B.
 DR Pfam: PF01376; Enterotoxin_B; 1.
 DR PRINTS: PR00772; ENTEROTOXINB.
 DR ProDom: PD012805; Enterotoxin_B; 1.
 DR Enterotoxin; Signal; 3D-structure.
 FT SIGNAL 1 21
 FT CHAIN 22 124 HEAT-LABILE ENTEROTOXIN B CHAIN.
 FT DISULFID 30 107
 FT VARIANT 6 6 F -> C (IN ISOLATE H10407).
 FT VARIANT 20 20 H -> Y (IN ISOLATE H10407).
 FT VARIANT 34 34 H -> R (IN ISOLATE H10407).
 SQ SEQUENCE 124 AA; 14027 MW; E9F7F7C7B9D3BC47 CRC64;

Query Match 100.0%; Score 62; DB 1; Length 124;
 Best local similarity 100.0%; Pred. No. 6e-05;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 VEVPGSQHDSQ 12
 Db 71 VEVPGSQHDSQ 82

RESULT 3
 ELPB_ECOLI
 ID ELPB_ECOLI STANDARD; PRT; 124 AA.

AC P32890; P13768; P01557;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)
 DE Heat-labile enterotoxin B chain precursor (LT-B, porcine) (LTP-B).
 GN ELPB OR LTPB.
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Escherichia.
 ON NCBI_TaxID=562;
 RX [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ISOLATE P307;
 RX MEDLINE=81074965; PubMed=7003397;
 RA Dallas W.S., Falkow S.;
 RT "Amino acid sequence homology between cholera toxin and Escherichia
 RT coli heat-labile toxin.";
 RL Nature 288:499-501(1980).
 RN [2]
 RP REVISIONS TO 28 AND 64.
 RC STRAIN-ISOLATE P307;
 RX MEDLINE=85156481; PubMed=3884513;
 RA Leong J., Vinal A.C., Dallas W.S.;
 RT "Nucleotide sequence comparison between heat-labile toxin B-subunit
 RT cistrons from Escherichia coli of human and porcine origin.";
 RL Infect. Immun. 48:73-77(1985).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ISOLATE PCG86;
 RX MEDLINE=87137303; PubMed=3546273;
 RA Yamamoto T., Gojobori T., Yokota T.;
 RT "Evolutionary origin of pathogenic determinants in enterotoxigenic
 RT Escherichia coli and Vibrio cholerae O1.";
 RL J. Bacteriol. 169:1352-1357(1987).
 RN [4]
 RP SEQUENCE OF 1-22 FROM N.A.
 RX MEDLINE=87280041; PubMed=3301830;
 RA Ibrahim I., Gentz R.;
 RT "A functional interaction between the signal peptide and the
 RT translation apparatus is detected by the use of a single point
 RT mutation which blocks translocation across mammalian endoplasmic
 RT reticulum.";
 RL J. Biol. Chem. 262:10189-10194(1987).
 RN [5]
 RP X-RAY CRYSTALLOGRAPHY (1.95 ANGSTROMS).
 RX MEDLINE=93240541; PubMed=8478941;
 RA Sikma T.K., van Zanten B.A.M., Dauter Z., Hol W.G.J.;
 RT "Refined structure of Escherichia coli heat-labile enterotoxin, a
 RT close relative of cholera toxin.";
 RL J. Mol. Biol. 230:890-918(1993).
 RN [6]
 RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS).
 RX MEDLINE=91238966; PubMed=2034287;
 RA Sikma T.K., Pronk S.E., Kalk K.H., Wieringa E.S., van Zanten B.A.M.,
 RA Witolt B., Hol W.G.J.;
 RT "Crystal structure of a cholera toxin-related heat-labile enterotoxin
 RT from E. coli.";
 RL Nature 351:371-377(1991).
 RN [7]
 RP DISCUSSION OF SEQUENCE.
 RX MEDLINE=95349400; PubMed=7623669;
 RA Domenighini M., Plaza M., Jobling M.G., Holmes R.K., Rappuoli R.;
 RT "Identification of errors among database sequence entries and
 RT comparison of correct amino acid sequences for the heat-labile
 RT enterotoxins of Escherichia coli and Vibrio cholerae.";
 RL Mol. Microbiol. 15:1165-1167(1995).
 CC -1- FUNCTION: THE BIOLOGICAL ACTIVITY OF THE TOXIN IS PRODUCED BY
 CC THE A CHAIN, WHICH ACTIVATES INTRACELLULAR ADENYL CYCLASE.
 CC -1- SUBUNIT: HETEROHEXAMER OF ONE A CHAIN AND OF FIVE B CHAINS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M17873; AAA98065.1; -
 DR EMBL: M15363; AAA24792.1; -
 DR EMBL: M17101; AAA23973.1; -
 DR PIR: A01820; QLECB.
 DR PIR: B26946; QLECB.
 DR PDB: 1LTA; 31-JAN-94.
 DR PDB: 1LTB; 31-JAN-94.
 DR PDB: 1LTG; 15-SEP-95.
 DR PDB: 1LTI; 17-AUG-96.
 DR PDB: 1LTS; 31-JAN-94.
 DR PDB: 1LTT; 31-JAN-94.
 DR PDB: 1LTK; 07-JUL-97.
 DR PDB: 1LTH; 16-JUN-97.
 DR PDB: 1LTF; 03-DEC-97.
 DR PDB: 1LTH; 03-DEC-97.
 DR PDB: 1LTH; 20-APR-95.
 DR InterPro: IPR001835; Enterotoxin_B.
 DR Pfam: PF01376; Enterotoxin_B. 1.
 DR PRINTS: PR00772; Enterotoxin_B.
 DR ProDom: PD012805; Enterotoxin_B. 1.
 DR Enterotoxin; Signal; 3D-structure.
 KW SIGNAL
 FT CHAIN 1 21
 FT DISULFID 22 124 HEAT-LABILE ENTEROTOXIN B CHAIN.
 FT HELIX 26 30
 FT TURN 31 32
 FT STRAND 36 43
 FT STRAND 47 51
 FT STRAND 54 55
 FT STRAND 58 62
 FT STRAND 64 65
 FT STRAND 68 71
 FT TURN 76 77
 FT TURN 80 98
 FT HELIX 99 100
 FT TURN 103 109
 FT STRAND 115 123
 FT STRAND 124 AA; 1413 MW; 6DB7DE58395EA70D CRC64;
 SQ SEQUENCE

Query Match 100.0%; Score 62; DB 1; Length 124;
 Best Local Similarity 100.0%; Pred. No. 6e-05;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VEPGSOHDSQ 12
 DB 71 VEPGSOHDSQ 82

RESULT 4
 DMB_MOUTH STANDARD; PRT; 674 AA.
 AC P27989;
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Carbon monoxide dehydrogenase beta subunit (EC 1.2.99.2) (CODH)
 OS Moorella thermoacetica (Clostridium thermoacetum).
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Moorella group;
 CC Moorella.
 NCBI_TaxID=1525;
 RN NCBI_TaxID=1525;
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92084676; PubMed=1748656;
 RA Morton T.A., Rungtist J.A., Ragsdale S.W., Shanmugasundaram T.,
 RA Wood H.G., Ljungdahl L.G.;
 RT "The primary structure of the subunits of carbon monoxide
 RT dehydrogenase/acetyl-CoA synthase from Clostridium thermoacetum.";

RL J. Biol. Chem. 266:23824-23828(1991).
 CC -I- FUNCTION: CATALYZES THE INTERCONVERSION OF CO AND CO2 AND THE
 CC SYNTHESIS OF ACETYL-COENZYME A FROM THE METHYLATED CORRIDOID/IRON
 CC SULEFUR PROTEIN, CO, AND COENZYME A.
 CC -I- CATALYTIC ACTIVITY: CO + H(2)O + acceptor = CO(2) + reduced
 CC acceptor.
 CC -I- COFACTOR: NICKEL.
 CC -I- SUBUNIT: TETRAMER OF TWO ALPHA AND TWO BETA CHAINS.
 CC -I- SIMILARITY: TO R. RUBRUM CARBON MONOXIDE DEHYDROGENASE.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M62727; AAA23228.1; -
 DR PIR: A41670; A41670.
 KW Oxidoreductase; Nickel; Iron-sulfur; Electron transport.
 FT METAL 68 68 IRON-SULEFUR (BY SIMILARITY).
 FT METAL 71 71 IRON-SULEFUR (BY SIMILARITY).
 FT METAL 90 90 IRON-SULEFUR (BY SIMILARITY).
 SQ SEQUENCE 674 AA; 72924 MW; 54BA3D816C25F9FC CRC64;

Query Match 61.3%; Score 38; DB 1; Length 674;
 Best Local Similarity 54.5%; Pred. No. 16;
 Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 2 EVPGSOHDSQ 12
 DB 379 KIPGAYHIDYQ 389

RESULT 5
 AROC_YEAST STANDARD; PRT; 376 AA.
 AC P28777;
 DT 01-DEC-1992 (Rel. 24, Created)
 DT 01-DEC-1992 (Rel. 24, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Chorismate synthase (EC 4.6.1.4) (5'-enolpyruvylshikimate-3-phosphate
 DE phospholase).
 GN ARO2 OR YGL148W.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 NCBI_TaxID=4932;
 RN NCBI_TaxID=4932;
 RP SEQUENCE FROM N.A.
 RC STRAIN=S288C;
 RX MEDLINE=92114793; PubMed=1837329;
 RA Jones D.G.L., Reusser U., Braus G.H.;
 RT "Molecular cloning, characterization and analysis of the regulation
 RT of the ARO2 gene, encoding chorismate synthase, of Saccharomyces
 RT cerevisiae.";
 RL Mol. Microbiol. 5:2143-2152(1991).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S288C / FY1769;
 RX MEDLINE=97197983; PubMed=9046099;
 RA Voet M., Defoor E., Verhasselt P., Riles L., Robben J., Volckaert G.;
 RT "The sequence of a nearly unclonable 22.8 kb segment on the left arm
 RT chromosome VII from Saccharomyces cerevisiae reveals ARO2, RPL9A,
 RT TIF1, MRF1 genes and six new open reading frames.";
 RL yeast 13:177-182(1997).
 CC -I- CATALYTIC ACTIVITY: 5-O-(1-carboxyvinyl)-3-phosphoshikimate =
 CC chorismate + phosphate.
 CC -I- COFACTOR: REDUCED FLAVIN.
 CC -I- PATHWAY: SEVENTH STEP IN THE BIOSYNTHESIS OF CHORISMATE WITHIN
 CC THE BIOSYNTHESIS OF AROMATIC AMINO ACIDS (THE SHIKIMATE PATHWAY).

```
CC -1- SUBUNIT: HOMOTETRAMER (BY SIMILARITY).
CC -1- INDUCTION: BY AMINO ACID STARVATION.
CC -1- SIMILARITY: BELONGS TO THE CHORISMATE SYNTHASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
```

DR EMBL: X60190: CAA42745.1; -;
DR EMBL: X99960: CAA68214.1; -;
DR EMBL: Z72670: CAA96860.1; -;
DR PIR: S17246: S17246.
DR SGD: S0003116: ARO2.
DR InterPro: IPR000453: Chorismate_synth.
DR Pfam: PF01264: Chorismate_synth; 1.
DR ProDom: PD002941: Chorismate_synth; 1.
DR PROSITE: PS00787: CHORISMATE_SYNTHASE_1; 1.
DR PROSITE: PS00788: CHORISMATE_SYNTHASE_2; 1.
DR PROSITE: PS00789: CHORISMATE_SYNTHASE_3; 1.
KM Lyase: Aromatic amino acid biosynthesis
SQ SEQUENCE 376 AA; 40838 MW; AF3AF65605B91B8E CRC64;

Query Match 59.7%; Score 37; DB 1; Length 376;
Best Local Similarity 70.0%; Pred. No. 13;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 VEPGSOHID 10
Db 264 VSPGSKHND 273

RESULT 6
YBSA_THEME STANDARD; PRT; 240 AA.
AC P58009;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein TM1158.1.
GN TM1158.1.
OS Thermotoga maritima.
OC Bacteria; Thermotogales; Thermotoga.
OX NCBI_TaxID=2336;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MSB8 / DSM 3109;
RX MEDLINE=99287316; PubMed=10360571;
RA Nelson K.E., Clayton R.A., Gill S.R., Gwin M.L., Dodson R.J.,
RA Haft D.H., Hickey E.K., Peterson J.D., Nelson W.C., Ketchum K.A.,
RA McDonald L., Utterback T.R., Malek J.A., Linher K.D., Garrett M.W.,
RA Stewart A.M., Cotton M.D., Pratt M.S., Phillips C.A., Richardson D.,
RA Heidelberg J., Sutton G.G., Fleischmann R.D., Eisen J.A., White O.,
RA Salzberg S.L., Smith H.O., Venter J.C., Fraser C.M.;
RT "Evidence for lateral gene transfer between Archaea and Bacteria from
RT genome sequence of Thermotoga maritima.";
RL Nature 399:323-329(1999).
RN [2]
RP IDENTIFICATION.
RA Medigue C., Bocs S.;
RT Unpublished observations (APR-2001).

```
CC -----
CC EMBL: AE001773; -; NOT_ANNOTATED_CDS.
CC TIGR: TM1158.1; -.
CC KW Hypothetical protein; Complete proteome.
CC SEQUENCE 240 AA; 27773 MW; 1BEF66C1C8BD2700 CRC64;
```

Query Match 58.1%; Score 36; DB 1; Length 240;
Best Local Similarity 50.0%; Pred. No. 13;
Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 3 VPGSOHIDSO 12
Db 226 VPGSEHLEK 235

RESULT 7
GUDH_BACSU STANDARD; PRT; 455 AA.
ID GUDH_BACSU
AC P4238;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Probable glucarate dehydratase (EC 4.2.1.40) (GDH) (GlucD).
GN GUDH.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group: Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=95219079; PubMed=7704254;
RA Ogawa K.-I., Akagawa E., Nakamura K., Yamane K.;
RT "Determination of a 21548 bp nucleotide sequence around the 24
RT degrees region of the Bacillus subtilis chromosome.";
RL Microbiology 141:269-275(1995).
CC -1- FUNCTION: CATALYZES THE DEHYDRATION OF GLUCARATE TO 5-KETO-4-
CC DEOXY-D-GLUCARATE (5-KDGLUC) (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: D-glucarate -> 5-dehydro-4-deoxy-D-glucarate +
CC H(2)O.
CC -1- PATHWAY: FIRST STEP IN GLUCARATE CATABOLISM.
CC -1- SIMILARITY: BELONGS TO THE MANDELATE RACEMASE / MUCONATE
CC LACTONIZING ENZYME FAMILY. GUDC SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: D30808: BAA06470.1; -;
DR EMBL: Z99105: CAB12043.1; -;
DR HSSP: P42206: 1BQG.
DR Subtilist; BG11161; gudD.
DR InterPro: IPR001354; MR_MLE.
DR Pfam: PF01188; MR_MLE; 1.
DR Pfam: PF02746; MR_MLE_N; 1.
KM Lyase: Complete proteome.
SQ SEQUENCE 455 AA; 50782 MW; 3238486007698C2A CRC64;

Query Match 58.1%; Score 36; DB 1; Length 455;
Best Local Similarity 75.0%; Pred. No. 26;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 EVPGSOH 9
Db 64 EVPGSEH 71

```
RESULT 8
ID BGAL_THEFT STANDARD: PRT: 743 AA.
AC P77989;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Beta-galactosidase (EC 3.2.1.23) (Lactase).
GN LAC2 OR LACA.
OS Thermoanaerobacter ethanolicus (Clostridium thermohydrosulfuricum).
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Thermoanaerobacter group; Thermoanaerobacter.
OX NCBI_TaxID=1757;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 33223 / 39E;
RA Zverlov V.;
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: Hydrolysis of terminal, non-reducing beta-D-
CC galactose residues in beta-D-galactosides.
CC -1- SIMILARITY: BELONGS TO FAMILY 2 OF GLYCOSYL HYDROLASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: Y08557; CAA69850.1; -
DR InterPro: IPR001649; Glyco_hydro_2.
DR Pfam: PF00703; Glyco_hydro_2; 1.
DR Pfam: PF02836; Glyco_hydro_2_C; 1.
DR Pfam: PF02837; Glyco_hydro_2_N; 1.
DR PROSITE: PS00719; GLYCOSYL_HYDROL_F2_1; 1.
DR PROSITE: PS00608; GLYCOSYL_HYDROL_F2_2; 1.
KM Hydrolyase; Glycosidase.
FT ACT_SITE 388 388 PROTON DONOR (BY SIMILARITY).
FT ACT_SITE 453 453 NUCLEOPHILE (BY SIMILARITY).
FT ACT_SITE 743 743 NOC11PFS17ES1DFC CRC64;
SQ SEQUENCE 743 AA; 85796 MW; FE011PFS17ES1DFC CRC64;

Query Match 58.1%; Score 36; DB 1; Length 743;
Best Local Similarity 54.5%; Pred. No. 44;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 EYVGSQHDSD 12
DB 346 EIRGMQHDDE 356

RESULT 9
ID ST14_HUMAN STANDARD: PRT: 855 AA.
AC Q9Y5Y6; Q9H3S0; Q9HCA3; Q9HS01; Q9HB36;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Suppressor of tumorigenicity 14 (EC 3.4.21.-) (Matrilinease) (Membrane-
DE type serine protease 1) (MT-SPI) (Prostasin) (Serine protease TADG-15)
DE (Tumor associated differentially-expressed gene-15 protein).
GN ST14 OR PRSS14 OR SMC19 OR TADG15.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=99303581; PubMed=10373424;
RA Lin C.Y., Anders J., Johnson M., Sang O.A., Dickson R.B.;
RT "Molecular cloning of cDNA for matrilinease, a matrix-degrading serine
RT protease with trypsin-like activity."
RT
SQ
```

PRINTS: PR00261; LDLRECEPTOR.
 DR SMART: SM00042; CUB: 2.
 DR SMART: SM00192; LDLa; 3.
 DR SMART: SM00020; TRYP_SPC; 1.
 DR PROSITE: PS01180; CUB: 2.
 DR PROSITE: PS01209; LDLRA_1; 2.
 DR PROSITE: PS50068; LDLRA_2; 4.
 DR PROSITE: PS50240; TRYPIN_DOM; 1.
 DR PROSITE: PS00134; TRYPIN_HIS; 1.
 DR PROSITE: PS00135; TRYPIN_SER; 1.
 DR Signal-anchor: Glycoprotein; Hydrolyase; Serine protease;
 KW Transmembrane; Repeat.
 FT DOMAIN 1 55 CYTOPLASMIC (POTENTIAL)
 FT TRANSSEM 56 76 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
 FT TRANSSEM 56 76 (POTENTIAL).
 FT DOMAIN 77 855 EXTRACELLULAR (POTENTIAL).
 FT DOMAIN 214 334 CUB 1.
 FT DOMAIN 340 447 CUB 2.
 FT DOMAIN 452 487 LDL-RECEPTOR CLASS A 1.
 FT DOMAIN 487 524 LDL-RECEPTOR CLASS A 2.
 FT DOMAIN 524 560 LDL-RECEPTOR CLASS A 3.
 FT DOMAIN 566 603 LDL-RECEPTOR CLASS A 4.
 FT DOMAIN 615 854 SERINE PROTEASE.
 FT ACT_SITE 656 656 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 711 711 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 805 805 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT CARBOHYD 109 109 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 302 302 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 485 485 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 772 772 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CONFLICT 327 329 FEA -> GTR (IN REF. 5).
 FT CONFLICT 381 381 R -> S (IN REF. 4).
 FT CONFLICT 674 674 A -> V (IN REF. 3).
 FT SEQUENCE 855 AA; 94769 MW; 26143132C01F99C9 CRC64;

Query Match 58.1%; Score 36; DB 1; Length 855;
 Best Local Similarity 55.6%; Pred. NO. 52;
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 VEPGSOHI 9
 Db 370 IEVNNNOHV 378
 RESULT 10
 ID GLS3_YEAST STANDARD; PRT: 1785 AA.
 AC 004952;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DE Putative 1,3-beta-glucan synthase component (EC 2.4.1.34) (1,3-beta-D-glucan-UDP glucosyltransferase).
 GN YMR306W OR YMR952.08.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycos.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S288C / AB972;
 RA Connor R., Churcher C.M., Barrell B.G., Rajandream M.A., Walsh S.V.;
 RL Submitted (APR-1995) to the EMBL/GenBank/DBJ databases.
 CC -1- CATALYTIC ACTIVITY: UDP-glucose + ((1,3)-beta-D-glucosyl)(N) = UDP + ((1,3)-beta-D-glucosyl)(N+1).
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 CC -1- SIMILARITY: STRONG TO GLS1 AND GLS2.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

DR EMBL: Z49212; CAAB9139.1; -
 DR SGD: S0004923; YMR306W.
 DR InterPro: IPR003440; Glucan_synthase.
 DR Pfam: PF02364; Glucan_synthase; 1.
 KW Hypothetical protein; Transmembrane; Transferase; Glycosyltransferase.
 FT TRANSSEM 337 357 POTENTIAL.
 FT TRANSSEM 375 395 POTENTIAL.
 FT TRANSSEM 415 435 POTENTIAL.
 FT TRANSSEM 444 464 POTENTIAL.
 FT TRANSSEM 508 528 POTENTIAL.
 FT TRANSSEM 547 567 POTENTIAL.
 FT TRANSSEM 572 592 POTENTIAL.
 FT TRANSSEM 712 732 POTENTIAL.
 FT TRANSSEM 1215 1235 POTENTIAL.
 FT TRANSSEM 1268 1288 POTENTIAL.
 FT TRANSSEM 1303 1323 POTENTIAL.
 FT TRANSSEM 1370 1390 POTENTIAL.
 FT TRANSSEM 1394 1414 POTENTIAL.
 FT TRANSSEM 1475 1495 POTENTIAL.
 FT TRANSSEM 1514 1534 POTENTIAL.
 FT TRANSSEM 1549 1569 POTENTIAL.
 FT TRANSSEM 1585 1605 POTENTIAL.
 FT TRANSSEM 1615 1675 POTENTIAL.
 FT TRANSSEM 1733 1733 POTENTIAL.
 SQ SEQUENCE 1785 AA; 207482 MW; 3475446DA46C6120 CRC64;

Query Match 58.1%; Score 36; DB 1; Length 1785;
 Best Local Similarity 54.5%; Pred. NO. 1,2e+02;
 Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 2 EVPGSOHDSQ 12
 Db 403 EWPGQHLSR 413
 RESULT 11
 ID PHEB_PSESP STANDARD; PRT: 302 AA.
 AC P31019;
 DT 01-JUL-1993 (Rel. 26, Created)
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Catechol 1,2-dioxygenase (EC 1.13.11.1).
 GN PHEB.
 OS Pseudomonas sp. (strain EST1001).
 OC Plasmid pEST1412.
 OC Bacteria; Proteobacteria.
 OX NCBI_TaxID=306;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91192610; PubMed=2013408;
 RA Kivisaar M., Kasak L., Nurk A.;
 RT "Sequence of the plasmid-encoded catechol 1,2-dioxygenase-expressing gene, pheB, of phenol-degrading Pseudomonas sp. strain EST1001.";
 RL Gene 98:15-20(1991).
 CC -1- CATALYTIC ACTIVITY: Catechol + O(2) = cis,cis-muconate.
 CC -1- COFACTOR: FERRIC ION.
 CC -1- PATHWAY: PHENOL DEGRADATION.
 CC -1- SIMILARITY: BELONGS TO THE INTRADIOL RING-CLEAVAGE DIOXYGENASE FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

```

CC -----
DR EMBL: M57500; AAC64900.1; -
DR PIR: JN0143; JN0143.
DR HSSP: P00437; 3PCD.
DR InterPro: IPR000627; Dioxigenase.
DR Pfam: PF00775; Dioxigenase; 1.
DR PROSITE: PS00083; INTRADIOL_DIOXYGENAS; 1.
DR Aromatic hydrocarbons catabolism; Oxidoreductase; Dioxigenase; Iron;
KW PLASmid.
FT METAL 164 164 IRON (BY SIMILARITY).
FT METAL 198 198 IRON (BY SIMILARITY).
FT METAL 222 222 IRON (BY SIMILARITY).
FT METAL 224 224 IRON (BY SIMILARITY).
SQ SEQUENCE 302 AA; 33362 MW; A86F17B68D1EAC3A CRC64;

Query Match
Best Local Similarity 41.7%; Score 35; DB 1; Length 302;
Matches 5; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

```

```

Qy 1 VEPGSOHIDSQ 12
Db 227 ISAPGHQHLTQ 238

```

```

RESULT 12
ID YP68_CAEEL STANDARD; PRT; 313 AA.
AC 009217;
DT 01-NOV-1997 (Rel. 35, Created)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical 37.0 kDa protein B0495.8 in chromosome II.
GN B0495.8.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA Kirsten J.;
RL Submitted (FEB-1995) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: SOME, TO YEAST YD1087C AND S.POMBE SPEC16A11.13.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U21317; AAA62527.1; -
DR WormRep: B0495.8; CE01766.
KW Hypothetical protein.
SQ SEQUENCE 313 AA; 36977 MW; 000D2327621BPED0 CRC64;

```

```

Query Match
Best Local Similarity 45.5%; Score 35; DB 1; Length 313;
Matches 5; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

```

```

Qy 2 EYVPSQHIDSQ 12
Db 11 QLMGSHVDNK 21

```

```

RESULT 13
ID YM18_PSEAE STANDARD; PRT; 371 AA.
AC 001609; G911P9;
DT 01-FEB-1995 (Rel. 31, Created)

```

```

DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein PA2218.
GN PA2218.
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxID=287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 15692 / PA01 / H103;
RX MEDLINE=93051258; PubMed=1427017;
RA Huang H., Stenmel R.J., Bellido F., Rawling E., Hancock R.E.W.;
RT "Analysis of two gene regions involved in the expression of the
RT imipenem-specific, outer membrane porin protein OprD of Pseudomonas
RT aeruginosa."
RT FEMS Microbiol. Lett. 76:267-274(1992).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 15692 / PA01;
RX MEDLINE=20437337; PubMed=10984043;
RA Stover C.K., Pham X.-O.T., Erwin A.L., Mizoguchi S.D., Warrenner P.,
RA Hickey M.J., Brinkman F.S.L., Huftagle W.O., Kowalik D.J., Lagrou M.,
RA Garber R.L., Goltzy L., Tolentino E., Westbrook-Wadman S., Yan Y.,
RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Laidig K., Lim R.M.,
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;
RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an
RT opportunistic pathogen."
RT Nature 406:959-964(2000).
RL Nature 406:959-964(2000).
CC -1- SIMILARITY: TO E.COLI YC9Y.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: Z14064; CAA78447.1; -
DR EMBL: AF004648; AAG05606.1; ALT_INIT.
DR PIR: S23861; S23861.
KW Hypothetical protein; Complete proteome.
FT CONFLICT 16 R -> P (IN REF. 1).
FT CONFLICT 73 R -> P (IN REF. 1).
FT CONFLICT 261 A -> G (IN REF. 1).
FT SEQUENCE 371 AA; 40840 MW; D7EB0CCACG5A7CF6 CRC64;

```

```

Query Match
Best Local Similarity 62.5%; Score 35; DB 1; Length 371;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

```

```

Qy 3 VPGASHVD 10
Db 343 VPGASHVD 350

```

```

RESULT 14
ID TYRO_ASPOR STANDARD; PRT; 539 AA.
AC 000234;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Tyrosinase (EC 1.14.18.1) (Monophenol monooxygenase).
GN MELO.
OS Aspergillus oryzae.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
OX NCBI_TaxID=5062;
RN [1]

```

RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RC STRAIN-ATCC 22788 / RIB 128;
 RX MEDLINE=95200965; PubMed-7893753;
 RA Fujita Y., Uraga Y., Ichishima E.;
 RT "Molecular cloning and nucleotide sequence of the protyrosinase gene,
 RT melo, from Aspergillus oryzae and expression of the gene in yeast
 RT cells.";
 RL Blochim. Biophys. Acta 1261:151-154(1995).
 CC -1- FUNCTION: THIS IS A COPPER-CONTAINING OXIDASE THAT FUNCTIONS IN
 CC THE FORMATION OF PIGMENTS SUCH AS MELANINS AND OTHER POLYPHENOLIC
 CC COMPOUNDS.
 CC -1- CATALYTIC ACTIVITY: L-tyrosine + L-DOPA + O(2) = L-DOPA +
 CC DOPAquinone + H(2)O.
 CC -1- COFACTOR: BINDS TWO COPPER IONS.
 CC -1- ENZYME REGULATION: ACTIVATED BY ACIDIFYING TREATMENT AT PH 3.0.
 CC -1- SUBUNIT: HOMOTETRAMER.
 CC -1- PTM: THE N-TERMINAL IS BLOCKED.
 CC -1- SIMILARITY: BELONGS TO THE TYROSINASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: D37929; BAA07149.1; -
 CC InterPro: IPR002227; Tyrosinase.
 CC Pfam: PF00264; Tyrosinase; 2.
 CC PRINTS: PR00092; TYROSINASE.
 CC PROSITE: PS00497; TYROSINASE_1; 1.
 CC PROSITE: PS00498; TYROSINASE_2; 1.
 CC DR PROSITE: PS00498; TYROSINASE_2; 1.
 CC KW Melanin biosynthesis; Oxidoreductase; Monooxygenase; Copper.
 CC FT METAL 63 63 COPPER A (BY SIMILARITY).
 CC FT METAL 84 84 COPPER A (BY SIMILARITY).
 CC FT METAL 93 93 COPPER A (BY SIMILARITY).
 CC FT METAL 290 290 COPPER B (BY SIMILARITY).
 CC FT METAL 294 294 COPPER B (BY SIMILARITY).
 CC FT METAL 333 333 COPPER B (BY SIMILARITY).
 CC FT THIOETH 82 84 BY SIMILARITY.
 CC SO SEQUENCE 539 AA; 60604 MW; CD2ECD702A018E15 CRC64;

Query Match 56.5%; Score 35; DB 1; Length 539;
 Best Local Similarity 54.5%; Pred. No. 48;
 Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 EYVPSQHDSQ 12
 | | | | |
 Db 360 EYVPSQHDSQ 370

RESULT 15
 INV1_MAIZE STANDARD; PRT; 670 AA.
 AC P49175;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Beta-fructofuranosidase 1 precursor (EC 3.2.1.26) (Sucrose-6-phosphate
 DE hydrolase 1) (invertase 1).
 GN IVR1.
 OS Zea mays (Maize).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
 OC Panicoideae; Andropogoneae; Zea.
 OC NCBI_TaxID=4577;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. B73;
 RX MEDLINE=95357417; PubMed=7630946;
 RX Xu J., Pemberton G.H., Almira E.C., McCarty D.R., Koch K.E.;

RT "The Ivr 1 gene for invertase in maize.";
 RL Plant Physiol. 108:1293-1294(1995).
 CC -1- CATALYTIC ACTIVITY: Hydrolysis of terminal non-reducing beta-D-
 CC fructofuranoside residues in beta-D-fructofuranosides.
 CC -1- SUBCELLULAR LOCATION: Vacuolar.
 CC -1- SIMILARITY: BELONGS TO FAMILY 32 OF GLYCOSYL HYDROLASES.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: U16123; AAA83439.1; -
 CC DR MaizeDB: 86037; -
 CC DR InterPro: IPR001362; Glyco_hydro_32.
 CC Pfam: PF00251; Glyco_hydro_32; 1.
 CC PROSITE: PS00609; GLYCOSYL_HYDROL_F32; 1.
 CC KW Hydrolase; Glycosidase; Glycoprotein; Zymogen; Signal;
 CC Multi-gene family.
 CC FT SIGNAL 1 2 POTENTIAL.
 CC FT PROPEP 2 2 POTENTIAL.
 CC FT CHAIN 1 670 BETA-FRUCTOFURANOSIDASE 1.
 CC FT ACT_SITE 139 139 BY SIMILARITY.
 CC FT CARBOHYD 165 165 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC FT CARBOHYD 275 275 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC FT CARBOHYD 518 518 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC FT CARBOHYD 595 595 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC FT CARBOHYD 639 639 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC SO SEQUENCE 670 AA; 71932 MW; DED0989C7E6A8E0 CRC64;

Query Match 56.5%; Score 35; DB 1; Length 670;
 Best Local Similarity 58.3%; Pred. No. 62;
 Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 VEYVPSQHDSQ 12
 | | | | |
 Db 61 VTYVPSQHDSQ 72

Search completed: July 3, 2002, 09:04:42
 Job time: 582 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 3, 2002, 09:06:02 ; Search time 74.17 Seconds
(without alignments)
27.989 Million cell updates/sec

Title: US-09-786-648-3

Perfect score: 62
Sequence: 1 VEVPSQHDSQ 12

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

1: SPREMBL_19:*
2: sp_archaea:*
3: sp_bacteria:*
4: sp_fungi:*
5: sp_human:*
6: sp_invertebrate:*
7: sp_mammal:*
8: sp_mhc:*
9: sp_organelle:*
10: sp_phage:*
11: sp_plant:*
12: sp_rodent:*
13: sp_virus:*
14: sp_vertebrate:*
15: sp_unclassified:*
16: sp_rvirus:*
17: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	62	100.0	103	2	Q9R646
2	62	100.0	124	2	Q57193
3	62	100.0	124	2	Q9R646
4	62	100.0	124	2	Q9R646
5	62	100.0	124	2	Q9R646
6	62	100.0	124	2	Q9R646
7	62	100.0	124	2	Q9R646
8	62	100.0	124	2	Q9R646
9	62	100.0	124	2	Q9R646
10	62	100.0	124	2	Q9R646
11	62	100.0	124	2	Q9R646
12	62	100.0	124	2	Q9R646
13	62	100.0	124	2	Q9R646
14	62	100.0	124	2	Q9R646
15	62	100.0	124	2	Q9R646
16	62	100.0	124	2	Q9R646

17	39	62.9	260	10	Q9MA16	Q9MA16 arabidopsis
18	39	62.9	272	4	Q9UG04	Q9UG04 homo sapien
19	39	62.9	353	16	Q9RVJ7	Q9RVJ7 deinococcus
20	39	62.9	412	4	Q9Y2B3	Q9Y2B3 homo sapien
21	39	62.9	428	10	Q9LGM2	Q9LGM2 oryza sativ
22	39	62.9	1076	10	Q94I82	Q94I82 oryza sativ
23	39	62.9	1242	13	Q90Y57	Q90Y57 brachydanio
24	38.5	62.1	565	10	Q22511	Q22511 vitis vinif
25	37	59.7	779	4	Q96HT2	Q96HT2 homo sapien
26	37	59.7	331	5	Q18391	Q18391 drosophila
27	37	59.7	354	3	Q14347	Q14347 schizosacch
28	37	59.7	367	10	Q80418	Q80418 nicotiana t
29	37	59.7	423	10	Q9C531	Q9C531 arabidopsis
30	37	59.7	457	10	Q949W1	Q949W1 arabidopsis
31	37	59.7	508	4	Q9987	Q9987 homo sapien
32	37	59.7	550	10	Q04636	Q04636 arabidopsis
33	37	59.7	899	3	Q96W69	Q96W69 candida alb
34	37	59.7	899	3	Q96W68	Q96W68 candida alb
35	37	59.7	1328	10	Q9AU17	Q9AU17 lycopersico
36	36	58.1	373	2	Q9XD79	Q9XD79 streptomyce
37	36	58.1	382	5	Q9V7M7	Q9V7M7 drosophila
38	36	58.1	389	5	Q9U0W5	Q9U0W5 leishmania
39	36	58.1	527	12	Q9E226	Q9E226 helicoverpa
40	36	58.1	528	12	Q99G79	Q99G79 heliocoverp
41	36	58.1	551	10	Q9S079	Q9S079 pinus taeda
42	36	58.1	745	2	Q93IM0	Q93IM0 thermoanaer
43	36	58.1	755	16	Q92XF7	Q92XF7 thizobium m
44	36	58.1	832	12	Q9D0H5	Q9D0H5 hepatitis b
45	36	58.1	832	12	Q9D0H1	Q9D0H1 hepatitis b

ALIGNMENTS

RESULT 1	
Q9R646	PRELIMINARY: PRT: 103 AA.
AC Q9R646:	
DT 01-MAY-2000 (TREMBLrel. 13, Created)	
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)	
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)	
DE CHOLERA-LIKE ENTEROTOXIN B SUBUNIT.	
OS Vibrio cholerae.	
OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.	
OX NCBI_TaxID=666;	
RN [1]	
RP SEQUENCE.	
RX MEDLINE=95303036; PubMed=7783690;	
RA Nakashima K., Eguchi Y., Nakasone N.;	
RT "Characterization of an enterotoxin produced by Vibrio cholerae	
RT O139."	
RL Microbiol. Immunol. 39:87-94(1995).	
DR HSSP; P01556; 1XTC.	
DR InterPro: IPR001835; Enterotoxin_B.	
DR Pfam: PF01376; Enterotoxin_B. 1.	
DR PRINTS: PR00772; ENTEROTOXINB.	
DR PRODOM: PD012805; Enterotoxin_B. 1.	
SQ SEQUENCE 103 AA; 11645 MW; 992A05C077F2F70E CRC64;	
Query Match	100.0%; Score 62; DB 2; Length 103;
Best Local Similarity	100.0%; Pred. NO. 0.00019;
Matches 12; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
OY 1 VEVPSQHDSQ 12	
DB 50 VEVPSQHDSQ 61	
RESULT 2	
Q57193	PRELIMINARY: PRT: 124 AA.
ID Q57193:	
AC Q57193:	

DT 01-NOV-1996 (TREMblrel. 01, Created)
DT 01-NOV-1996 (TREMblrel. 01, last sequence update)
DT 01-JUN-2001 (TREMblrel. 17, last annotation update)
DE CHOLERA TOXIN B PROTEIN (CTB) PRECURSOR (CTB).
GN CTXB.
OS *Vibrio cholerae*.
OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; *Vibrio*.
OX NCBI_TaxID=666;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CLASSICAL STRAIN 569B;
RX MEDLINE=91355224; PubMed=1883840;
RA Dams E., De Wolf M., Dierick W.;
RT "Nucleotide sequence analysis of the CT operon of the *Vibrio cholerae*
RL classical strain 569B.";
RL Biochim. Biophys. Acta 1090:139-141(1991).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CLASSICAL BIOTYPE 569B;
RA Shi C., Cao C., Zhang J., Ma Q.;
RL Chin. Biochem. J. 9:395-399(1993).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=CLASSICAL BIOTYPE 569B;
RA Xu L.;
RL Submitted (APR-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL: X58785; CAA41591.1; -;
DR EMBL: U25679; AAC34728.1; -;
DR EMBL: A00931; CAA00098.1; -;
DR HSSP: P01556; 2CHB.
DR InterPro: IPR001835; Enterotoxin_B.
DR Pfam: PF01376; Enterotoxin_B.1.
DR PRINTS: PR00772; ENTEROTOXINB.
DR ProDom: PD012805; Enterotoxin_B.1.
KW Signal.
FT SIGNAL 1 21 POTENTIAL.
FT CHAIN 22 124 CHOLERA TOXIN B PROTEIN (CTB).
SQ SEQUENCE 124 AA; 13919 MW; D6BF83FEF7924EA3 CRC64;

Query Match 100.0%; Score 62; DB 2; Length 124;
Best Local Similarity 100.0%; Pred. No. 0.00023;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VEVPGSQHIDSQ 12
Db 71 VEVPGSQHIDSQ 82
RESULT 3
ID 09RP15 PRELIMINARY; PRT; 124 AA.
AC 09RP15;
DT 01-MAY-2000 (TREMblrel. 13, Created)
DT 01-MAY-2000 (TREMblrel. 13, last sequence update)
DT 01-JUN-2001 (TREMblrel. 17, last annotation update)
DE CHOLERA ENTEROTOXIN B-SUBUNIT.
GN CTXB.
OS *Vibrio cholerae*.
OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; *Vibrio*.
OX NCBI_TaxID=666;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=KNH002;
RA Shin H.J., Park Y.C., Kim Y.C.;
RT "Cloning and nucleotide sequence analysis of the virulence gene
RT cassette from *Vibrio cholerae* KNH002 isolated in Korea.";
RL Misailurrahag Hoji 35:205-210(1999).
DR EMBL: AF175708; AAD51360.1; -;
DR HSSP: P01556; 2CHB.
DR InterPro: IPR001835; Enterotoxin_B.
DR Pfam: PF01376; Enterotoxin_B.1.
DR PRINTS: PR00772; ENTEROTOXINB.

DR ProDom: PD012805; Enterotoxin_B.1.
SQ SEQUENCE 124 AA; 13905 MW; 23BF83FEF793E5B9 CRC64;

Query Match 100.0%; Score 62; DB 2; Length 124;
Best Local Similarity 100.0%; Pred. No. 0.00023;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VEVPGSQHIDSQ 12
Db 71 VEVPGSQHIDSQ 82

RESULT 4
ID 056635 PRELIMINARY; PRT; 124 AA.
AC 056635;
DT 01-NOV-1996 (TREMblrel. 01, Created)
DT 01-NOV-1996 (TREMblrel. 01, last sequence update)
DT 01-JUN-2001 (TREMblrel. 17, last annotation update)
DE CHOLERA TOXIN PRECURSOR.
GN CTXB.
OS *Vibrio cholerae*.
OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; *Vibrio*.
OX NCBI_TaxID=666;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=57;
RA Yamamoto K., Do V.G., Xu M., Iida T., Miwatani T., Albert M.J.,
RA Honda T.;
RT "Comparison of cholera toxin genes (ctxAB) of non-O1 *Vibrio cholerae*
RT strains 854 (O139-bengal) and 57 (O37) from two outbreaks.";
RL Submitted (APR-1994) to the EMBL/GenBank/DBJ databases.
DR EMBL: D30052; BAA06289.1; -;
DR HSSP: P01556; 2CHB.
DR InterPro: IPR001835; Enterotoxin_B.
DR Pfam: PF01376; Enterotoxin_B.1.
DR PRINTS: PR00772; ENTEROTOXINB.
DR ProDom: PD012805; Enterotoxin_B.1.
KW Signal.
FT SIGNAL 1 21 POTENTIAL.
FT SEQUENCE 124 AA; 13871 MW; 3f87B2F297953179 CRC64;

Query Match 100.0%; Score 62; DB 2; Length 124;
Best Local Similarity 100.0%; Pred. No. 0.00023;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VEVPGSQHIDSQ 12
Db 71 VEVPGSQHIDSQ 82

RESULT 5
ID 093V32 PRELIMINARY; PRT; 124 AA.
AC 093V32;
DT 01-DEC-2001 (TREMblrel. 19, Created)
DT 01-DEC-2001 (TREMblrel. 19, last sequence update)
DT 01-DEC-2001 (TREMblrel. 19, last annotation update)
DE HEAT-LABILE ENTEROTOXIN B SUBUNIT.
GN LTH B SUBUNIT.
OS *Escherichia coli*.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae; *Escherichia*.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1032 (ENTEROTOXIGENIC);
RA Komase K.;
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.

RC STRAIN-1032 (ENTEROTOXIGENIC);
 RX MEDLINE=95091056; PubMed=7998417;
 RA Tamura S., Asanuma H., Tomita T., Komase K., Kawahara K., Danbara H.,
 RA Hattori N., Watanabe K., Suzuki Y., Nagamine T., Aizawa C., Oya A.,
 RA Kurata T.;
 RT "Escherichia coli heat-labile enterotoxin B subunits supplemented with
 RT a trace amount of the holotoxin as an adjuvant for nasal influenza
 RT vaccine.";
 RL Vaccine 12:1083-1089(1994).
 DR EMBL: AB011677; BAZ25726.1;
 SQ SEQUENCE 124 AA; 14028 MW; 5346BD3BB32354C2 CRC64;

Query Match 100.0%; Score 62; DB 2; Length 124;
 Best Local Similarity 100.0%; Pred. No. 0.00023;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VEPGSOHIDSQ 12
 Db 71 VEPGSOHIDSQ 82

RESULT 6
 O94M01 PRELIMINARY; PRT; 124 AA.
 AC O94M01;
 DT 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE CTXB.
 GN CTXB.
 OS Vibrio phage CTX.
 CC Viruses; ssDNA viruses; Inoviridae; Inovirus.
 OX NCBI_TaxID=141904;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Bhattacharyya T., Nandy R.K., Nair G.B.;
 RT "The entire core region of the ctx-phi (ctx-prophage) in VCE 232, an
 RT environmental strain of V. cholerae.";
 RL Submitted (Aug-2001) to the EMBL/Genbank/DBJ databases.
 DR EMBL: AF414369; AAL09682.1;
 SQ SEQUENCE 124 AA; 13884 MW; 8D4C251591B16891 CRC64;

Query Match 100.0%; Score 62; DB 9; Length 124;
 Best Local Similarity 100.0%; Pred. No. 0.00023;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 VEPGSOHIDSQ 12
 Db 71 VEPGSOHIDSQ 82

RESULT 7
 O9NKM5 PRELIMINARY; PRT; 395 AA.
 AC O9NKM5;
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
 DE HYPOHETICAL 43.6 KDA PROTEIN.
 GN BG.DS01514.3.
 OS Drosophila melanogaster (Fruit fly).
 CC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 CC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 CC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA STRAIN-Y, CN BW SP;
 RX MEDLINE=99403001; PubMed=10471707;
 RA Ashburner M., Misra S., Roote J., Lewis S.E., Blazey R., Davis T.,
 RA Doyle C., Galle R., George R., Harris N., Hartnell G., Harvey D.,

RA Hong L., Houston K., Hoskins R., Johnson G., Martin C., Moshrefi A.,
 RA Palazzolo M., Reese M.G., Spradling A., Tsang G., Wan K., Whitehead K.,
 RA Celniker S., Rubin G.M.;
 RT "An exploration of the sequence of a 2.9-Mb region of the genome of
 RT Drosophila melanogaster: the Adh region.";
 RL Genetics 153:179-219(1999).
 RN [2]

RP SEQUENCE FROM N.A.
 RC STRAIN-Y, CN BW SP;
 RA Celniker S.E., Agbayani A., Arcalua T.T., Baxter E., Blazey R.G.,
 RA Butenoff C., Champe M., Chavez C., Chew M., Ciesiolka L., Doyle C.M.,
 RA Farfan D.E., Galle R., George R.A., Harris N.L., Hoskins R.A.,
 RA Houston K.A., Hummasti S.R., Karra K., Kearney L., Kim E., Lee B.,
 RA Lewis S., Li P., Lomoton M.A., Mazda P., Moshrefi A.R., Moshrefi M.,
 RA Nixon K., Paclab J.M., Park S., Pfeiffer B., Poon L., Sequeira A.,
 RA Sehl H., Snir E., Svirska R.R., Wan K.H., Weinburg T., Zhang R.,
 RA Zieran L.L., Rubin G.M.;
 RL Submitted (MAR-2000) to the EMBL/Genbank/DBJ databases.
 DR EMBL: AE003408; AAF44847.1;
 DR FlyBase: FBgn0028907; BG.DS01514.3.
 KW Hypothetical protein.
 SQ SEQUENCE 395 AA; 43561 MW; AE4F1CC4AD3DA73 CRC64;

Query Match 69.4%; Score 43; DB 5; Length 395;
 Best Local Similarity 72.7%; Pred. No. 4.2;
 Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 OY 1 VEPGSOHIDS 11
 Db 330 VSPGSHIDA 340

RESULT 8
 O947W6 PRELIMINARY; PRT; 802 AA.
 AC O947W6;
 DT 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE PUTATIVE TNP-LIKE TRANSPOSABLE ELEMENT.
 GN OSJNBA0029P06.9.
 OS Oryza sativa (Rice)
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 CC Euphorbiaceae; Oryzae; Oryza.
 OX NCBI_TaxID=4530;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Wing R.A., Yu Y., Soderlund C., Chen M., Kim H.-R., Rambo T.,
 RA Saeki C., Henry D., Oates R., Simmons J.;
 RT "Rice Genomic Sequence.";
 RL Submitted (SEP-2001) to the EMBL/Genbank/DBJ databases.
 DR EMBL: AC090873; AAK98653.1;
 SQ SEQUENCE 802 AA; 92934 MW; 014DE54BA1CD108 CRC64;

Query Match 67.7%; Score 42; DB 10; Length 802;
 Best Local Similarity 77.8%; Pred. No. 14;
 Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 4 PGSOHIDSQ 12
 Db 516 PGSOHIESE 524

RESULT 9
 O94HM6 PRELIMINARY; PRT; 1109 AA.
 AC O94HM6;
 DT 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)

DE PUTATIVE TRANSPOSABLE ELEMENT.

OS Oryza sativa (Rice).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Eriarthroideae; Oryzaceae; Oryza.
 OX NCBI_TaxID=4530;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NIPPONBARE;
 RA Spiegel L., de la Bastide M., Nascimento L., Kirchoff K., King L.,
 Preston R., Vil M.D., Baker J., Bell M., Zutavern T., Santos L.,
 RA Miller B., Kuit K., Rodriguez S., Cummins D.M., Balija V., Shah R.,
 RA Bahret A., Bal H., O'Shaughnessy A., Dedhia N., McCombie W.R.;
 RT Genomic Sequence for Oryza sativa, Nipponbare strain, clone
 OSJNB0089D15, from Chromosome 10, complete sequence."
 RL Submitted (Aug-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NIPPONBARE;
 RA McCombie W.R.;
 RL Submitted (Aug-2000) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NIPPONBARE;
 RA McCombie W.R.;
 RL Submitted (May-2001) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NIPPONBARE;
 RA Palmer L.E., Spiegel L., de la Bastide M., Nascimento L., Kirchoff K.,
 RA King L., Preston R., Vil M.D., Baker J., Bell M., Zutavern T.,
 RA Santos L., Miller B., Kuit K., Rodriguez S., Cummins D.M., Balija V.,
 RA Shah R., Bahret A., Bal H., O'Shaughnessy A., Dedhia N.,
 RA McCombie W.R.;
 RL Submitted (Aug-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AC079844; AAK92611.1; -;
 SQ SEQUENCE 1109 AA; 127116 MW; 5C026FEF70EDB855 CRC64;

Query Match

Best Local Similarity 67.7%; Score 42; DB 10; Length 1109;
 Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 4 PGSQHDSQ 12
 Db 526 PGSQHISE 534

RESULT 10

ID 094EB4 PRELIMINARY; PRT; 1352 AA.

AC 094EB4;
 DT 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
 DE P0002B05.25 PROTEIN.
 GN P0002B05.25.
 OS Oryza sativa (Rice).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Eriarthroideae; Oryzaceae; Oryza.
 OX NCBI_TaxID=4530;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. NIPPONBARE;
 RA Sasaki T., Matsumoto T., Yamamoto K.;
 RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC
 clone:P0002B05."
 RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AP001141; BAB63505.1; -;
 SQ SEQUENCE 1352 AA; 154782 MW; DFF265A2B971B8F8 CRC64;

Query Match

Best Local Similarity 67.7%; Score 42; DB 10; Length 1352;
 Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 4 PGSQHDSQ 12
 Db 463 PGSQHISE 471

RESULT 11

ID 0948C7 PRELIMINARY; PRT; 2535 AA.

AC 0948C7;
 DT 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
 DE PUTATIVE TNP-LIKE TRANSPOSABLE ELEMENT.
 GN OSJNB0034A02.4.
 OS Oryza sativa (Rice).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Eriarthroideae; Oryzaceae; Oryza.
 OX NCBI_TaxID=4530;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Wang R.A., Yu Y., Soderlund C., Chen M., Kim H.-R., Rambo T.,
 RA Sackl C., Henry D., Oates R., Simmons J.;
 RT "Rice Genomic Sequence."
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AC079852; AAL25171.1; -;
 SQ SEQUENCE 2535 AA; 289895 MW; DD280569D154168B CRC64;

Query Match

Best Local Similarity 67.7%; Score 42; DB 10; Length 2535;
 Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 4 PGSQHDSQ 12
 Db 526 PGSQHISE 534

RESULT 12

ID 086582 PRELIMINARY; PRT; 242 AA.

AC 086582;
 DT 01-NOV-1998 (TREMBlrel. 08, Created)
 DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
 DE HYPOTHEICAL. 26.0 KDA PROTEIN.
 GN SC2A11.21C.
 OS Streptomyces coelicolor.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
 OX NCBI_TaxID=1902;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=A3(2);
 RA Murphy L., Harris D.;
 RL Submitted (Aug-1998) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=A3(2);
 RA Parhill J., Bartell B.G., Rajandream M.A.;
 RL Submitted (Aug-1998) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=A3(2);
 RX MEDLINE=97000351; PubMed=8843436;
 RA Redenbach M., Kieser H.M., Denapalte D., Elchner A., Cullum J.,
 RA Kinsahl H., Hopwood D.A.;
 RT "A set of ordered cosmid and a detailed genetic and physical map for
 the 8 Mb Streptomyces coelicolor A3(2) chromosome.";

RL Mol. Microbiol. 21:77-96(1996).
 DR EMBL: AL031184; CAA20190.1; -;
 DR InterPro: IPR002502; Amidase_2.
 DR Pfam: PF01510; Amidase_2; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 242 AA; 26018 MW; 6DD10FF18A2EC544 CRC64;

Query Match
 Best Local Similarity 64.5%; Score 40; DB 2; Length 242;
 Pred. No. 9.3;
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Oy 1 VEPGSOHID 10
 |||||: ||
 Db 204 VEPGTDHTD 213

RESULT 13
 ID 031335 PRELIMINARY; PRT; 254 AA.
 AC 031335;
 DT 01-JAN-1998 (TREMblrel. 05, Created)
 DT 01-JAN-1998 (TREMblrel. 05, Last sequence update)
 DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
 DE AMP-DEPENDENT NUCLEASE SUBUNIT A (FRAGMENT).
 CN ADDA.
 OS Bacillus cereus.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group;
 OC Bacillus/Staphylococcus group; Bacillus.
 OX NCBI_TaxID=1396;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 10987;
 RA MEDLINE=99231848; PubMed=10217496;
 RA Ostad O.A., Hegna I., Lindbaeck T., Rishovd A.L., Kolsto A.B.;
 RT "Genome organisation is not conserved between Bacillus cereus and
 RT Bacillus subtilis";
 RL Microbiology 145:621-631(1999).
 DR EMBL: Y11217; CAA72103.1; -;
 DR InterPro: IPR000212; Vrd-helicase.
 DR Pfam: PF00580; Vrd-helicase; 1.
 FT NON_TER 1 1
 FT 254 254
 SQ SEQUENCE 254 AA; 29212 MW; 24392E1133BD992A CRC64;

Query Match
 Best Local Similarity 64.5%; Score 40; DB 2; Length 254;
 Pred. No. 9.8;
 Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Oy 1 VEPGSOHID 12
 :: ||||| |
 Db 35 IDEPGSHIRKO 46

RESULT 14
 ID 09NP06 PRELIMINARY; PRT; 192 AA.
 AC 09NP06;
 DT 01-OCT-2000 (TREMblrel. 15, Created)
 DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)
 DT 01-OCT-2000 (TREMblrel. 15, Last annotation update)
 DE LfLp, LCAT-LIKE LYOPHOSPHOLIPASE (FRAGMENT).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Pluvinet R., Estivill X., Escarceller M., Sumoy L.;
 RL submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.

RA Auffray C., Ansgore W., Ballabio A., Estivill X., Gibson K.,
 RA Lehrach H., Poustka A., Lundberg J.;
 RT "The European IMAGE consortium for integrated molecular analysis of
 RT human gene transcripts."
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AL389957; CAB97531.1; -;
 FT NON_TER 1 1
 SQ SEQUENCE 192 AA; 21609 MW; 04A7AE8CB344F213 CRC64;

Query Match
 Best Local Similarity 62.9%; Score 39; DB 4; Length 192;
 Pred. No. 11;
 Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Oy 2 EYPGSOHID 10
 |||||: ||
 Db 166 ELPGESEHIE 174

RESULT 15
 ID 09RJ28 PRELIMINARY; PRT; 201 AA.
 AC 09RJ28;
 DT 01-MAY-2000 (TREMblrel. 13, Created)
 DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)
 DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
 DE PUTATIVE AMIDASE.
 GN SCG11A.03.
 OS Streptomyces coelicolor.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
 OX NCBI_TaxID=1902;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=A3(2);
 RA Oliver K., Harris D.;
 RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=A3(2);
 RA Redenbach M., Kieser H.M., Denapate D., Eichner A., Cullum J.,
 RA Kinashl H., Hopwood D.A.;
 RT "A set of ordered cosmids and a detailed genetic and physical map for
 RT the 8 Mb Streptomyces coelicolor A3(2) chromosome."
 RL Mol. Microbiol. 21:77-96(1996).
 DR EMBL: AL133210; CAB61584.1; -;
 DR InterPro: IPR002502; Amidase_2.
 DR Pfam: PF01510; Amidase_2; 1.
 SQ SEQUENCE 201 AA; 22749 MW; B8EFA77E06A20468 CRC64;

Query Match
 Best Local Similarity 62.9%; Score 39; DB 2; Length 201;
 Pred. No. 12;
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 2 EYPGSOHID 10
 |||||: ||
 Db 176 EYPGSDHTD 184

Search completed: July 3, 2002, 09:06:03
 Job time: 617 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 3, 2002, 08:54:57 ; Search time 94.52 Seconds

(without alignments)
24.678 Million cell updates/sec

Title: US-09-786-648-4
Perfect score: 106
Sequence: 1 GATFOVEVPGSQHIDSQKAI 21

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Maximum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

A.Geneseq_032802:*

- 1: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1980.DAT.*
- 2: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1981.DAT.*
- 3: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1982.DAT.*
- 4: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1983.DAT.*
- 5: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1984.DAT.*
- 6: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1985.DAT.*
- 7: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1986.DAT.*
- 8: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1987.DAT.*
- 9: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1988.DAT.*
- 10: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1989.DAT.*
- 11: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1990.DAT.*
- 12: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1991.DAT.*
- 13: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1992.DAT.*
- 14: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1993.DAT.*
- 15: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1994.DAT.*
- 16: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1995.DAT.*
- 17: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1996.DAT.*
- 18: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1997.DAT.*
- 19: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1998.DAT.*
- 20: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1999.DAT.*
- 21: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA2000.DAT.*
- 22: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	106	100.0	21	AAV87462	Cholera toxin B/en
2	106	100.0	103	AAW04857	Synthetic cholera
3	106	100.0	103	AAK94939	Heat labile entero
4	106	100.0	103	AAW06606	Cholera toxin B su
5	106	100.0	103	AAW06607	Cholera toxin B su
6	106	100.0	103	AAW80808	Amino acid sequenc
7	106	100.0	103	AAW62367	V. cholera cholera
8	106	100.0	103	AAW62370	V. cholera cholera
9	106	100.0	105	AAW62369	V. cholera cholera
10	106	100.0	118	AAW04163	Cholera Toxin B-su
11	106	100.0	123	AAW62374	E. coli LTB protei

12	106	100.0	124	10	AAW93561	B subunit of the h
13	106	100.0	124	17	AAW06605	Cholera toxin B su
14	106	100.0	124	19	AAW59770	Amino acid sequenc
15	106	100.0	124	21	AAW96652	Plant-optimized E.
16	106	100.0	124	21	AAW59992	Cholera toxin B su
17	106	100.0	124	22	AAW62359	V. cholera strain
18	106	100.0	124	22	AAW62363	V. cholera cholera
19	106	100.0	124	22	AAW62368	V. cholera cholera
20	106	100.0	124	22	AAW62376	E. coli LTB protei
21	106	100.0	124	22	AAW62378	E. coli LTB protei
22	106	100.0	125	22	AAW62372	E. coli LTB protei
23	106	100.0	131	11	AAW04825	LTB-CTB fusion pro
24	106	100.0	138	15	AAW50227	Sequence of LT-B-M
25	106	100.0	170	20	AAW94082	LTB-CTP fusion pro
26	106	100.0	380	22	AAW00507	E. coli heat-labill
27	106	100.0	382	22	AAW00506	E. coli heat-labill
28	106	100.0	405	12	AAW11272	HSV-1 antigen/heat
29	103	97.2	103	6	AAW50340	Sequence of sub-un
30	101	95.3	21	21	AAW87463	E. coli heat labill
31	101	95.3	93	16	AAW72545	ADP-ribosylating t
32	101	95.3	93	20	AAW41816	Escherichia coli v
33	101	95.3	93	20	AAW95226	E. coli heat-labill
34	101	95.3	93	21	AAW68365	Heat labile toxin
35	101	95.3	93	22	AAW66239	E coli verotoxin-1
36	101	95.3	103	22	AAW62379	E. coli LTB protei
37	101	95.3	123	22	AAW62377	E. coli LTB protei
38	101	95.3	124	22	AAW62373	E. coli LTB protei
39	101	95.3	124	22	AAW62375	E. coli LTB protei
40	101	95.3	124	22	AAW62380	E. coli LTB protei
41	101	95.3	134	22	AAW73241	Recombinant exotox
42	101	95.3	142	22	AAW73242	Recombinant exotox
43	101	95.3	155	22	AAW73243	Recombinant exotox
44	101	95.3	153	22	AAW73244	Recombinant exotox
45	101	95.3	371	20	AAW01300	Labile toxin (LT-B

ALIGNMENTS

RESULT 1

ID AAV87462 standard; peptide; 21 AA.

XX AAV87462:

XX 03-JUL-2000 (first entry)

DE Cholera toxin B/enterotoxin B-derived peptide, SEQ ID NO:4.

XX

KW Cholera toxin subunit B; CtxB; heat labile enterotoxin subunit B; EtxB;

KW beta-4-alpha-2 loop; GM-1 ganglioside receptor; immunomodulation;

KW adjuvant; immune disorder; diarrhoea.

XX

OS Vibrio cholerae.

OS Escherichia coli.

PN WO200014114-A1.

PD 16-MAR-2000.

PF 07-SEP-1999; 99WC-GH02970.

PR 07-SEP-1998; 98GB-0019484.

PA (UYBR-) UNIV BRISTOL.

PI Williams NA, Hirst TR;

DR WPT; 2000-256943/22.

XX

PT Derivatives of Escherichia coli heat labile enterotoxins useful as

PT immunomodulators and for treating diarrhoea and which do not bind the

PT glycolipid receptor GM-1 -

XX Disclosure; Page 15; 62pp; English.

PS
XX The invention relates to peptide fragments of the *Escherichia coli* heat
CC labile enterotoxin (Etx) and its closely related homologue, cholera
CC toxin (Ctx) from *Vibrio cholerae* which do not bind to the ubiquitous
CC GM-1 ganglioside receptors on host cell surfaces. Both Etx and Ctx are
CC composed of one A subunit and five identical B subunits. The A subunit
CC is responsible for toxicity, possessing adenosine diphosphate (ADP)-
CC ribosyltransferase activity, while the B subunits (EtxB and CtxB)
CC facilitate the entry of subunit A into the host cell via the binding and
CC cross-linking of GM-1 receptors. Although GM-1 binding is responsible
CC for some of the effects of Etx and Ctx, it has been found that certain
CC effects of the toxins, such as immunomodulation, are not mediated
CC through GM-1 binding. The peptides of the invention are fragments of the
CC beta-4- alpha-2 loop of EtxB and/or CtxB, exerting the same effects as
CC normal EtxB and CtxB subunits, except that they do not bind or cross link
CC GM-1. They may be used in medicine as an immunomodulator or adjuvant.
CC They may also be used as an inhibitor for toxin-induced diarrhoea.
CC Therefore, the peptides may be used in the production of a composition
CC for treating, preventing and/or modulating a disease associated with an
CC immune disorder and/or toxin-induced diarrhoea. Sequences AAY87460-Y87463
CC represent preferred peptides of the invention, AAY87460 being
CC particularly preferred.

XX
XX Sequence 21 AA:

SQ

Query Match 100.0%; Score 106; DB 21; Length 21;
Best Local Similarity 100.0%; Pred. No. 5.1e-11;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GATFOVEVPSGSHDSQKKAI 21
|||
Db 1 gatfvevpsghidsqskai 21

RESULT 2
AAM04857
ID AAM04857 standard; Protein; 103 AA.

XX
XX AAM04857;

AC

XX 21-FEB-1997 (first entry)

DT
XX
XX
XX Synthetic cholera toxin B subunit.

DE
XX
XX Bordetella pertussis; whooping cough; recombinant construct;
KM cholera toxin B subunit; enzyme; antigen; immunogen; allergen;
KW enzyme inhibitor; hormone; lymphokine; immunoglobulin; toxin;
KM structural protein; receptor; heterologous gene; leader; promoter.

XX
XX Synthetic.

OS
XX
XX W09626282-A1.

PN
XX
XX 29-AUG-1996.

PD
XX
XX 23-FEB-1996; 96WO-CA00107.

PF
XX
XX 23-FEB-1995; 95US-0393334.

PR
XX
XX (CONN-) CONNAUGHT LAB LTD.

PA
XX
XX Klein MH, Loosmore SM, Yacoub RK, Zealey GR;
PI
XX
XX WPI; 1996-425088/42.
DR
XX
XX N-PSDB; AAT38038.

PT
XX
XX Recombinant constructs for expressing and opt. secreting proteins in
PT Bordetella - comprise Bordetella promoter coupled to non-Bordetella,
PT esp. cholera B toxin, gene or coupled to non-Bordetella leader and
PT gene of interest

XX Example 1; Figure 1; 61pp; English.

PS
XX
XX Recombinant constructs comprising a promoter functional in
CC Bordetella operatively linked to a heterologous gene or a non-
CC Bordetella leader sequence for secretion of a gene product which may
CC or may not be of Bordetella origin, can be used for the expression
CC in Bordetella of enzymes, antigens, immunogens, allergens, enzyme
CC inhibitors, hormones, lymphokines, immunoglobulins or their
CC fragments, toxins, mammalian proteins, structural proteins or
CC receptors. The Bordetella strains are particularly engineered to
CC express the cholera toxin B subunit (this sequence). The promoters
CC used in the constructs are selected from the Bordetella pertussis
CC tox, fha promoters or the high molecular weight (hmw) outer membrane
CC promoter of non typable *Haemophilus influenzae*; leaders used in the
CC constructs are selected from the cholera toxin B leader (CTB-L), the
CC pertussis toxin subunit S1 leader (S1-L) and the pertussin peractin
CC leader (PRN-L); and genes used in the constructs are selected from a
CC novel synthetic cholera toxin B gene (ctb) and the hmw1 and hmw2
CC genes of *Haemophilus influenzae*.

XX
XX Sequence 103 AA:

SQ

Query Match 100.0%; Score 106; DB 17; Length 103;
Best Local Similarity 100.0%; Pred. No. 3.2e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GATFOVEVPSGSHDSQKKAI 21
|||
Db 45 gatfvevpsghidsqskai 65

RESULT 3
AAR94939
ID AAR94939 standard; Protein; 103 AA.

XX
XX AAR94939;

AC

XX 31-OCT-1996 (first entry)

DT
XX
XX
XX Heat labile enterotoxin B subunit (LT-B) *E.coli*.

DE
XX
XX Toxin; subunit; vaccine; transgenic plant; immunogen; antigen;
KW adjuvant; immunisation.

XX
XX
XX *Escherichia coli*.

OS
XX
XX W09612801-A1.

PN
XX
XX 02-MAY-1996.

PD
XX
XX 24-OCT-1995; 95WO-US13376.

PF
XX
XX 24-OCT-1994; 94US-0328716.

PR
XX
XX (TULA) TULANE EDUCATIONAL FUND.
PA (TEXA) UNIV TEXAS A & M SYSTEM.

XX
XX
XX Arntzen CJ, Clements JD, Haq TA, Mason HS;
PI
XX
XX WPI; 1996-230602/23.
DR
XX
XX N-PSDB; AAT18799, AAT18800.

PT
XX
XX Transgenic plants contg. *E. coli* heat labile enterotoxin subunits
PT used as oral vaccines for animals which consume the plant

PS
XX
XX Disclosure; Page 100-101; 130pp; English.

CC A transgenic plant comprising or expressing a DNA sequence encoding
CC an immunogenic agent can be used as an oral vaccine for animals.
CC The vaccine is administered by the oral consumption of the plant and
CC provides the first known functional method for immunising animals

CC using transgenic plants, where the plants express bacterial antigens
CC that act as both immunogens and adjuvants. The method provides an
CC inexpensive production and delivery system for such antigens to
CC animals. This is the LT-B Escherichia coli toxin subunit and its
CC coding sequence was used in the construction of such a transgenic
CC plant. The immunogenic agent preferably comprises the LT-B or CT-B
CC (cholera toxin B subunit) or optionally LT-A or CT-A.

XX Sequence 103 AA;

Query Match 100.0%; Score 106; DB 17; Length 103;
Best Local Similarity 100.0%; Pred. No. 3.2e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GATFQVEVPSGSHDSQKAI 21
Db 45 gatfvevpsgshdsqkai 65

RESULT 4

AAW06606 standard; Protein; 103 AA.

XX AAW06606;

DT 06-AUG-1997 (first entry)

XX Cholera toxin B subunit/heat labile enterotoxin B subunit hybrid.

XX Heat labile; tolerant; stable; hybrid; cholera toxin B subunit; CTB;

XX heat labile enterotoxin B subunit; LTB; vaccine; immunisation;

XX immunogen; antigen; broad spectrum; cross reactive; diarrhoea;

XX vomiting; food poisoning.

XX Chimeric - Vibrio cholerae.

OS Chimeric - Enterotoxigenic Escherichia coli.

XX key Location/Qualifiers

FT Misc-difference 1

FT /label= substitution

FT /note= "wild-type Thr replaced by Ala"

FT Misc-difference 94

FT /label= substitution

FT /note= "wild-type His replaced by Asn"

FT Misc-difference 95

FT /label= substitution

FT /note= "wild-type Ala replaced by Ser"

XX WO9634893-A1.

XX 07-NOV-1996.

XX 02-MAY-1996; 96WO-SE00570.

XX 05-MAY-1995; 95SE-0001682.

XX (HOLM/) HOLMGREN J.

XX (LEBE/) LEBENS M R.

XX Holmgren J, Lebens MR;

XX WPI; 1996-506108/50.

XX N-PSDB; AAT43576.

XX Heat-labile enterotoxin B sub-unit and cholera toxin B sub-unit

XX hybrid protein - opt. fused to immunogenic sequence for use in

XX vaccines against enterotoxin-induced illness

XX Claim 3; Page -; 32pp; English.

XX AAW06606 is a mature cholera toxin B subunit (CTB)/heat labile

XX enterotoxin B subunit (LTB) hybrid molecule. In mature CTB protein,

CC certain amino acids (aa) were replaced with corresponding aa from
CC heat-labile enterotoxin B subunit (LTB). The specific amino acid
CC substitutions impart LTB-specific epitope characteristics to
CC immunogenic mature CTB. The hybrid molecules have increased to
CC cross-reactivity and are suitable for a broad spectrum vaccine to
CC protect against enterotoxigenic illness. Immunogenic proteins
CC comprising the hybrid molecules can be used to treat, or in a
CC vaccine to prevent, enterotoxin-induced illness, e.g. diarrhoea
CC and vomiting, in humans and animals.

XX Note - This sequence does not appear in the specification, it is

XX a claimed mutant sequence of mature cholera toxin B subunit (see

XX AAW06605).

XX Sequence 103 AA;

Query Match 100.0%; Score 106; DB 17; Length 103;
Best Local Similarity 100.0%; Pred. No. 3.2e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GATFQVEVPSGSHDSQKAI 21
Db 45 gatfvevpsgshdsqkai 65

RESULT 5

AAW06607 standard; Protein; 103 AA.

XX AAW06607;

DT 06-AUG-1997 (first entry)

XX Cholera toxin B subunit/heat labile enterotoxin B subunit hybrid.

XX Heat labile; tolerant; stable; hybrid; cholera toxin B subunit; CTB;

XX heat labile enterotoxin B subunit; LTB; vaccine; immunisation;

XX immunogen; antigen; broad spectrum; cross reactive; diarrhoea;

XX vomiting; food poisoning.

XX Chimeric - Vibrio cholerae.

OS Chimeric - Enterotoxigenic Escherichia coli.

XX key Location/Qualifiers

FT Misc-difference 1..25

FT /label= substitution

FT /note= "the first 25 amino acids of mature

FT wild-type cholera toxin B subunit are

FT replaced with the first 25 amino acids

FT of mature enterotoxin B subunit"

XX WO9634893-A1.

XX 07-NOV-1996.

XX 02-MAY-1996; 96WO-SE00570.

XX 05-MAY-1995; 95SE-0001682.

XX (HOLM/) HOLMGREN J.

XX (LEBE/) LEBENS M R.

XX Holmgren J, Lebens MR;

XX WPI; 1996-506108/50.

XX N-PSDB; AAT43577.

XX Heat-labile enterotoxin B sub-unit and cholera toxin B sub-unit

XX hybrid protein - opt. fused to immunogenic sequence for use in

XX vaccines against enterotoxin-induced illness

XX Claim 4; Page -; 32pp; English.

AAW06607 is a mature cholera toxin B subunit (CTB)/heat labile enterotoxin B subunit (LTB) hybrid molecule. In mature CTB protein, certain amino acids (aa) were replaced with corresponding aa from heat-labile enterotoxin B subunit (LTB). The specific amino acid substitutions impart LTB-specific characteristics to immunogenic mature CTB. The hybrid molecules have increased cross-reactivity and are suitable for a broad spectrum vaccine to protect against enterotoxigenic illness. Immunogenic proteins comprising the hybrid molecules can be used to treat, or in a vaccine to prevent, enterotoxin-induced illness, e.g. diarrhoea and vomiting, in humans and animals.

Note - This sequence does not appear in the specification, it is a claimed mutant sequence of mature cholera toxin B subunit (see AAW06605).

Sequence 103 AA:

Query Match 100.0%; Score 106; DB 17; Length 103;
Best Local Similarity 100.0%; Pred. No. 3.2e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GATFOVEVPGSOHDSOKKAI 21
Db 45 gatfvevpgsqhdsqkai 65

RESULT 6
ID AAW80808 standard; protein; 103 AA.

AAW80808;
29-JAN-1999 (first entry)

Amino acid sequence of the wild type cholera toxin B subunit.

Cholera toxin B subunit; nontoxic subunit; adjuvant; coadministration;
antigen; bird; animal; mucosal; vaccine.

Vibrio cholerae.

WO9845324-A1.

15-OCT-1998.

03-APR-1998; 98WO-US06725.

04-APR-1997; 97US-0043410.

(KIYO/) KIYONO H.
(MCGH/) MCGHEE J R.
(TAKE/) TAKEDA Y.
(UABR-) UAB RES FOUND.
(YAMA/) YAMAMOTO S.

Kiyono H, Mcghee JR, Takeda Y, Yamamoto S;

WPI; 1998-594478/50.

New mutant cholera toxin selected from a group comprising nontoxic subunits/derivatives - effective as an adjuvant when coadministered with an antigen to birds and mammals

Disclosure: Fig 1B; 43pp; English.

This is the amino acid sequence of the cholera toxin B subunit used in the method of the invention involving the use of nontoxic subunits as an effective adjuvant in coadministration of an antigen to birds and animals. In addition to the use of the toxin as an mucosal adjuvant, it also provides a vaccine comprising the toxin, an immunogenic amount of an antigen, and a pharmaceutically acceptable carrier. The toxin can be used with single/multiple vaccines, and it

enables the possibility for commercial mucosal adjuvants for use in humans, since these are more effective and safer than vaccines administered subcutaneously.

Sequence 103 AA:

Query Match 100.0%; Score 106; DB 19; Length 103;
Best Local Similarity 100.0%; Pred. No. 3.2e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GATFOVEVPGSOHDSOKKAI 21
Db 45 gatfvevpgsqhdsqkai 65

RESULT 7

AAW62367
ID AAW62367 standard; protein; 103 AA.

AAW62367;

29-JUN-2001 (first entry)

V. cholera cholera toxin B (CTB) protein variant (GI 1421511).

AB5B subunit protein; mutation; recombinant; cholera toxin B; CTB;
immunogen; autoimmune disease; neuron demyelinating disease; rCTB;
infection; parasitic; immunosuppressive; antiarthritic; antirheumatic;
antidiabetic; neuroprotective; vaccine; variant.

Vibrio cholerae.

WO200127144-A2.

19-APR-2001.

05-OCT-2000; 2000WO-US27607.

08-OCT-1999; 99US-0158561.

(ACTI-) ACTIVE BIOTECH AB.

Handley HH, Haaparanta T, Ewalt KL;

WPI; 2001-281974/29.

Recombinant AB5B subunit protein comprising a mutation that alters the number residues available for chemical modification, useful for covalently linking to an immunogen or vaccine which can be used for treating autoimmune diseases -

Disclosure: Page 67-68; 78pp; English.

The invention relates to a recombinant AB5B subunit protein comprising at least one mutation, where the mutation alters the number of amino acid residues available for chemical modification as compared to a wild type AB5B subunit protein, and where the recombinant protein retains an effective target ligand binding affinity. A recombinant AB5B subunit protein such as cholera toxin B protein (CTB) can be specifically covalently linked at lysines or cysteines to an immunogen or vaccine. Recombinant mutant CTB (rCTB) can be used in the treatment of autoimmune diseases e.g. rheumatoid arthritis, encephalomyelitis (or other neuron demyelinating diseases) and diabetes. The rCTB or other B-subunits of the invention can also be used to induce tolerance to infection, e.g. parasitic infection. The present sequence represents a CTB variant protein, NCBI GenBank No. GI 1421511.

Sequence 103 AA:

Query Match 100.0%; Score 106; DB 22; Length 103;
Best Local Similarity 100.0%; Pred. No. 3.2e-10;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GATFOVEVPGSHIDSQKAI 21
|||||
DB 45 gatlqvevpgshidsqskai 65

RESULT 8

AAB62370
ID AAB62370 standard; Protein: 103 AA.

AC AAB62370;

DT 29-JUN-2001 (first entry)

DE V. cholera cholera toxin B (CTB) protein variant (GI 1421525).

KW AB5B subunit protein; mutation; recombinant; cholera toxin B; CTB;
immunogen; autoimmune disease; neuron demyelinating disease; rCTB;
infection; parasitic; immunosuppressive; antiarthritic; antineumatic;
antidiabetic; neuroprotective; vaccine; variant.

OS Vibrio cholerae.

PN WO200127144-A2.

PD 19-APR-2001.

PF .05-OCT-2000; 2000WO-US27607.

PR 08-OCT-1999; 99US-0158561.

PA (ACTI-) ACTIVE BIOTECH AB.

PI Handley HH, Haaparanta T, Ewalt KL;

PS WPI; 2001-281974/29.

PT Recombinant AB5B subunit protein comprising a mutation that alters the
number residues available for chemical modification, useful for
covalently linking to an immunogen or vaccine which can be used for
treating autoimmune diseases -

PS Disclosure: Page 69; 78pp; English.

CC The invention relates to a recombinant AB5B subunit protein comprising at
least one mutation, where the mutation alters the number of amino acid
residues available for chemical modification as compared to a wild type
CC AB5B subunit protein, and where the recombinant protein retains an
effective target ligand binding affinity. A recombinant AB5B subunit
protein such as cholera toxin B protein (CTB) can be specifically
covalently linked at lysines or cysteines to an immunogen or vaccine.
CC Recombinant mutant CTB (rCTB) can be used in the treatment of autoimmune
diseases e.g. rheumatoid arthritis, encephalomyelitis (or other neuron
demyelinating diseases) and diabetes. The rCTB or other B subunits of the
invention can also be used to induce tolerance to infection, e.g.
CC parasitic infection. The present sequence represents a CTB variant
protein, NCBI Genbank No. GI 1421525 (Ogawa 41 R35D).

CC Sequence 103 AA;

Query Match 100.0%; Score 106; DB 22; Length 103;
Best Local Similarity 100.0%; Pred. No. 3.2e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GATFOVEVPGSHIDSQKAI 21
|||||
DB 45 gatlqvevpgshidsqskai 65

RESULT 9

AAB62369

ID AAB62369 standard; Protein: 105 AA.

AC AAB62369;

DT 29-JUN-2001 (first entry)

DE V. cholera cholera toxin B (CTB) protein variant (GI 2781121).

KW AB5B subunit protein; mutation; recombinant; cholera toxin B; CTB;
immunogen; autoimmune disease; neuron demyelinating disease; rCTB;
infection; parasitic; immunosuppressive; antiarthritic; antineumatic;
antidiabetic; neuroprotective; vaccine; variant.

OS Vibrio cholerae.

PN WO200127144-A2.

PD 19-APR-2001.

PF 05-OCT-2000; 2000WO-US27607.

PR 08-OCT-1999; 99US-0158561.

PA (ACTI-) ACTIVE BIOTECH AB.

PI Handley HH, Haaparanta T, Ewalt KL;

PS WPI; 2001-281974/29.

PT Recombinant AB5B subunit protein comprising a mutation that alters the
number residues available for chemical modification, useful for
covalently linking to an immunogen or vaccine which can be used for
treating autoimmune diseases -

PS Disclosure: Page 68-69; 78pp; English.

CC The invention relates to a recombinant AB5B subunit protein comprising at
least one mutation, where the mutation alters the number of amino acid
residues available for chemical modification as compared to a wild type
CC AB5B subunit protein, and where the recombinant protein retains an
effective target ligand binding affinity. A recombinant AB5B subunit
protein such as cholera toxin B protein (CTB) can be specifically
covalently linked at lysines or cysteines to an immunogen or vaccine.
CC Recombinant mutant CTB (rCTB) can be used in the treatment of autoimmune
diseases e.g. rheumatoid arthritis, encephalomyelitis (or other neuron
demyelinating diseases) and diabetes. The rCTB or other B subunits of the
invention can also be used to induce tolerance to infection, e.g.
CC parasitic infection. The present sequence represents a CTB variant
protein, NCBI Genbank No. GI 2781121 (Ogawa 41).

CC Sequence 105 AA;

Query Match 100.0%; Score 106; DB 22; Length 105;
Best Local Similarity 100.0%; Pred. No. 3.3e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GATFOVEVPGSHIDSQKAI 21
|||||
DB 46 gatlqvevpgshidsqskai 66

RESULT 10

AAR04163
ID AAR04163 standard; protein: 118 AA.

AC AAR04163;

DT 10-SEP-1990 (first entry)

DE Cholera Toxin B-subunit.

KW cholera toxin B-subunit; hybrid protein; heterologous IGA active antigen.

XX synthetic.
OS
XX
FH Key Location/Qualifiers
FT misc_difference 18..18
FT /*label= His or Tyr
FT region 1..11
FT /*label= signal peptide
FT /*note= absent from mature protein
XX
XX WO3003437-A.
XX
XX PD 05-APR-1990.
XX
XX PF 27-SEP-1989; 89WO-0000495.
XX
XX PR 27-SEP-1988; 88FR-0012627.
XX
XX PA (UYLI-) L'UNIVERSITE DE L'ETAT A LIEGE.
XX
XX PI L'Hoir C, Renard A, Martial J;
XX
XX DR WPI: 1990-132273/17.
XX
XX DR N-NSDB: Q04046.
XX
XX PT New hybrid protein, useful in vaccines -
XX PT contains cholera toxin b subunit and heterologous Iga active
XX PT antigenic sequence.
XX
XX PS Disclosure; ; pp: French.
XX
XX CC Mature cholera toxin B-subunit is obtained when the signal peptide is
XX CC cleaved off. There is an Ochre codon at position 343-5; the sequence
XX CC downstream from it is part of a plasmid.
XX
XX SQ Sequence 118 AA;

Query Match 100.0%; Score 106; DB 11; Length 118;
Best Local Similarity 100.0%; Pred. No. 3.8e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GATFOVEVPGSGHIDSOKKAI 21
DB 56 gattfvevpgsqhidsqkai 76
RESULT 11
AAB62374
ID AAB62374 standard; Protein: 123 AA.
XX
XX AAB62374;
XX
XX DT 29-JUN-2001 (first entry)
XX
XX DE E. coli LTB protein variant (GI: 1395122).
XX
XX KW ABSB subunit protein; mutation; recombinant; cholera toxin B; CTB;
XX KW immunogen; autoimmune disease; neuron demyelinating disease; rCTB; LTB;
XX KW infection; parastitic; immunosuppressive; antiarthritic; antirheumatic;
XX KW antidiabetic; neuroprotective; vaccine; heat labile toxin B; variant.
XX
XX OS Escherichia coli.
XX
XX PN WO200127144-A2.
XX
XX PD 19-APR-2001.
XX
XX PF 05-OCT-2000; 2000WO-US27607.
XX
XX PR 08-OCT-1999; 99US-0158561.
XX
XX PA (ACTI-) ACTIVE BIOTECN AB.

XX Handley HH, Haaparanta T, Ewalt KL;
XX
XX PI WPI: 2001-281974/29.
XX
XX DR
XX
XX PT Recombinant ABSB subunit protein comprising a mutation that alters the
XX PT number residues available for chemical modification, useful for
XX PT covalently linking to an immunogen or vaccine which can be used for
XX PT treating autoimmune diseases -
XX
XX PS Disclosure; page 70; 78pp; English.
XX
XX
XX CC The invention relates to a recombinant ABSB subunit protein comprising at
XX CC least one mutation, where the mutation alters the number of amino acid
XX CC residues available for chemical modification as compared to a wild type
XX CC ABSB subunit protein, and where the recombinant protein retains an
XX CC effective target ligand binding affinity. A recombinant ABSB subunit
XX CC protein such as cholera toxin B protein (CTB) can be specifically
XX CC covalently linked at lysines or cysteines to an immunogen or vaccine.
XX CC Recombinant mutant CTB (rCTB) can be used in the treatment of autoimmune
XX CC diseases e.g. rheumatoid arthritis, encephalomyelitis (or other neuron
XX CC demyelinating diseases) and diabetes. The rCTB or other B subunits of the
XX CC invention can also be used to induce tolerance to infection, e.g.
XX CC parasitic infection. The present sequence represents a variant of the
XX CC E. coli heat labile toxin B (LTB) protein.
XX
XX SQ Sequence 123 AA;

Query Match 100.0%; Score 106; DB 22; Length 123;
Best Local Similarity 100.0%; Pred. No. 4e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GATFOVEVPGSGHIDSOKKAI 21
DB 66 gattfvevpgsqhidsqkai 86
RESULT 12
AAP93561
ID AAP93561 standard; protein: 124 AA.
XX
XX AAP93561;
XX
XX DT 06-JUN-1990 (first entry)
XX
XX DE B subunit of the heat-labile enterotoxin (LT-B) derived from E. coli.
XX
XX KW B subunit; heat-labile enterotoxin; LT-B; Escherichia coli; malaria;
XX KW circumsporozoite protein; fusion protein; live recombinant vaccine;
XX KW Salmonella; epitope.
XX
XX OS Escherichia coli.
XX
XX FH Key Location/Qualifiers
XX FH Peptide 1..22
XX FT /note="Signal peptide"
XX FT Protein 23..124
XX FT /note="Mature LT-B"
XX
XX PN WO8902924-A.
XX
XX PD 06-APR-1989.
XX
XX PF 30-SEP-1988; 88WO-US03376.
XX
XX PR 02-OCT-1987; 87US-0104735.
XX
XX PA (PRAK-) PRAXIS BIOLOGICS IN.
XX
XX PI Bray RN, Majarian WR, Pillai S, Hockmeyer WT;
XX
XX DR WPI: 1989-114399/15.

DR N-PSDB; AAN90747.
 XX
 XX Live recombinant vaccine for malaria -
 PT comprising attenuated entero-invasive bacterium contg. DNA
 PT encoding epitope of malaria parasite
 XX
 PS Fig 3; p. 3/17; 105bp; English.
 XX
 CC In the patent, the DNA encoding LT-B is expressed as part of a fusion
 CC protein with an epitope of a malaria parasite, eg Region I or Region II
 CC or a repeat region of circumsporozoite protein antigen (CS) (AAP3560)
 CC from Plasmodium berghei. Pref. the fusion gene is inserted into
 CC attenuated Salmonella enteritidis under the left promoter control of
 CC lambda. Such bacteria can multiply in the host without causing disease or
 CC disorder and express CS that will induce a protective immune response
 CC against malaria and can be used in vaccines. Such vaccines can be
 CC multivalent.
 CC
 XX
 SQ Sequence 124 AA;
 XX
 Query Match 100.0%; Score 106; DB 10; Length 124;
 Best Local Similarity 100.0%; Pred. No. 4e-10;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 GATFOVEVPGSQHTDSQKKA1 21
 ||||||||||||||||||||
 DB 66 gatfvevpgsqhdsqkka1 86
 RESULT 13
 AAM06605
 ID AAM06605 standard; Protein; 124 AA.
 XX
 AC AAM06605;
 XX
 DT 06-AUG-1997 (first entry)
 XX
 DE Cholera toxin B subunit, used for hybrid immunogenic toxin production.
 XX
 KW Heat labile; tolerant; stable; hybrid; cholera toxin B subunit; CTB;
 KW heat labile enterotoxin B subunit; LTb; vaccine; immunisation;
 KW immunogen; antigen; broad spectrum; cross reactive; diarrhoea;
 KW vomiting; food poisoning.
 XX
 OS Vibrio cholerae.
 XX
 FH Key Location/Qualifiers
 FT Peptide 1..21
 FT /label= sig_peptide
 FT Protein 22..124
 FT /label= mat_protein
 XX
 PN WO9634893-A1.
 XX
 PD 07-NOV-1996.
 XX
 PF 02-MAY-1996; 96WO-SE00570.
 XX
 PR 05-MAY-1995; 95SE-0001682.
 XX
 PA (HOLM/) HOLMGREN J.
 PA (LEBE/) LEBENS M R.
 XX
 PI Holmgren J, Lebens MR;
 XX
 DR WPI: 1996-506108/50.
 DR N-PSDB; AAT43575.
 XX
 XX Heat-labile enterotoxin B sub-unit and cholera toxin B sub-unit
 PT hybrid protein - opt. fused to immunogenic sequence for use in
 PT vaccines against enterotoxin-induced illness
 XX

PS Disclosure; Fig 1; 32pp; English.
 XX
 CC AAM06605 is the full (including the signal peptide) length cholera
 CC toxin B subunit (CTB), this sequence is described as unpublished in
 CC the specification. The mature CTB protein was used to create hybrid
 CC mutants, in which certain amino acids (aa) of CTB were replaced with
 CC corresponding aa from heat-labile enterotoxin B subunit (LTb), see
 CC AAM06606 and AAM06607. The specific amino acid substitutions impart
 CC LTb-specific epitope characteristics to immunogenic mature CTB. The
 CC hybrid molecules have increased cross-reactivity and are suitable
 CC for a broad spectrum vaccine to protect against enterotoxigenic
 CC illness. Immunogenic proteins comprising the hybrid molecules can be
 CC used to treat, or in a vaccine to prevent, enterotoxin-induced illness,
 CC e.g. diarrhoea and vomiting, in humans and animals.
 CC
 XX
 SQ Sequence 124 AA;
 XX
 Query Match 100.0%; Score 106; DB 17; Length 124;
 Best Local Similarity 100.0%; Pred. No. 4e-10;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 GATFOVEVPGSQHTDSQKKA1 21
 ||||||||||||||||||||
 DB 66 gatfvevpgsqhdsqkka1 86
 RESULT 14
 AAM59770
 ID AAM59770 standard; Protein; 124 AA.
 XX
 AC AAM59770;
 XX
 DT 12-OCT-1998 (first entry)
 XX
 DE Amino acid sequence of E. coli LTb.
 XX
 KW Beta-subunit of heat labile enterotoxin; LTb; fusion protein; vaccine;
 KW immunogen; antigen; inhibin; fertility; follicle stimulating hormone;
 KW FSH; sperm; ova; immune response.
 XX
 OS Escherichia coli.
 XX
 PN WO9821344-A1.
 XX
 PD 22-MAY-1998.
 XX
 PF 12-NOV-1997; 97WO-US20584.
 XX
 PR 12-NOV-1996; 96US-0747410.
 XX
 PA (UNMS) UNIV MICHIGAN STATE.
 XX
 PI Bagdasarian M, Ireland J;
 XX
 DR WPI: 1998-297947/26.
 DR N-PSDB; AAV41573.
 XX
 PT New nucleic acid encoding fusion of antigenic peptide and
 PT enterotoxin sub-unit - useful as vaccinating immunogen, particularly
 PT for increasing animal fertility by inducing antibodies against
 PT inhibin
 XX
 PS Disclosure; Fig 9; 56pp; English.
 XX
 CC This is the amino acid sequence of Escherichia coli beta-subunit of
 CC heat labile enterotoxin (LTb). It is used in the method of the
 CC invention to create fusion proteins which are useful as vaccinating
 CC immunogens. The fusion proteins are useful in vaccines, specifically
 CC where the antigenic peptide is an inhibin fragment for increasing the
 CC fertility of an animal (by increasing levels of follicle stimulating
 CC hormone (FSH) or production of sperm or ova), but more generally for
 CC inducing an immune response against the antigenic peptide. Vaccines

CC are particularly administered orally, e.g. fusion protein is expressed
 CC in edible plants or animals.
 XX
 SQ Sequence 124 AA:

Query Match 100.0%; Score 106; DB 19; Length 124;
 Best Local Similarity 100.0%; Pred. No. 4e-10;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GATFOVEVPGSQHDSQKKAI 21
 DB 66 gatfqvevpgsqhdsqkka 86

RESULT 15
 AAY96652
 ID AAY96652 standard: Protein; 124 AA.
 XX
 AC AAY96652;

DT 26-SEP-2000 (first entry)

DE Plant-optimized E. coli heat labile toxin B subunit.

XX Heat-labile toxin; LT-A; LT-B; mutant; transgenic plant; vaccine; oral;
 KW adjuvant; anti-bacterial.

XX Escherichia coli.
 OS Synthetic.

XX Key Location/Qualifiers
 FH Peptide 1..21
 FT /label= signal_peptide
 FT 22..124
 FT /label= mature_protein

XX WO200037609-A2.

XX 29-JUN-2000.

XX 22-DEC-1999; 99WO-US30747.

XX 22-DEC-1998; 98US-0113507.

XX (BOYC-) BOYCE THOMPSON INST PLANT RES.
 PA (MASO/) MASON H S.
 PA (ARNT/) ARNTZEN C J.

XX Mason HS, Arntzen CJ;

XX WPI: 2000-442653/38.
 DR N-PSDB: AAA51156.

XX New polynucleotides encoding LT-A or CT-A polypeptides for the
 PT transformation of plant cells, useful in immunogenic compositions to
 PT elicit immune responses in animals
 XX
 PS Claim 4; Fig 5A-B; 103pp; English.

XX This synthetic Escherichia coli heat-labile toxin (LT) B subunit (LT-B)
 CC is encoded by a plant-codon optimized cDNA. The cDNA sequence contains
 CC plant-preferred codons and eliminates sequence motifs associated with
 CC spurious mRNA processing. The second codon is changed from AAT encoding
 CC Asn to GTG encoding Val, in order to create a NcoI restriction site at
 CC the 5' end. Novel polynucleotides encode a mutant LT-A polypeptide or a
 CC mutant Vibrio cholerae cholera toxin (CT) A subunit (CT-A) polypeptide,
 CC which have reduced enzyme activity as compared to the wild-type LT-A or
 CC CT-A polypeptide and where at least one of the codons is altered to a
 CC plant preferred codon. The polynucleotide further comprises a nucleic
 CC acid sequence encoding LT B subunit (LT-B) or a CT B subunit (CT-B). The
 CC polynucleotides are useful for the transformation of plant cells for the
 CC production of transgenic plants to produce edible vaccines, especially

CC oral vaccines in transgenic plants for the prophylactic or therapeutic
 CC treatment against E. coli or V. cholerae. The mutant polypeptides are
 CC also useful as adjuvants.
 XX
 SQ Sequence 124 AA:

Query Match 100.0%; Score 106; DB 21; Length 124;
 Best Local Similarity 100.0%; Pred. No. 4e-10;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GATFOVEVPGSQHDSQKKAI 21
 DB 66 gatfqvevpgsqhdsqkka 86

Search completed: July 3, 2002, 08:54:57
 Job time: 287 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 3, 2002, 08:55:39 ; Search time 34.88 Seconds
(without alignments)
14.706 Million cell updates/sec

Title: US-09-786-648-4

Perfect score: 106

Sequence: 1 GATFOVEYVPGSQHDSOKKAI 21

Scoring table: BLOSUM62

Searched: Gapop 10.0 , Gapext 0.5

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

Issued_Patents_AA: *
1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep: *
2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep: *
3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep: *
4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep: *
5: /cgn2_6/ptodata/2/1aa/PCNUS.COMB.pep: *
6: /cgn2_6/ptodata/2/1aa/backfile1.pep: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	106	100.0	102	3	US-08-952-337-5
2	106	100.0	102	3	US-08-952-337-6
3	106	100.0	103	2	US-08-472-171-2
4	106	100.0	103	2	US-08-894-526-2
5	106	100.0	103	2	US-09-013-047-2
6	106	100.0	103	4	US-09-374-597-2
7	106	100.0	103	4	US-09-191-852-21
8	106	100.0	103	5	PCT-US95-13376-21
9	106	100.0	123	3	US-08-952-337-1
10	106	100.0	123	3	US-08-952-337-2
11	106	100.0	124	2	US-08-747-410-2
12	101	95.3	93	2	US-08-292-968-26
13	101	95.3	93	2	US-08-467-974-26
14	101	95.3	93	2	US-08-467-536-26
15	101	95.3	93	3	US-08-467-976-26
16	101	95.3	93	4	US-09-082-514-26
17	101	95.3	371	2	US-08-829-026A-6
18	94	88.7	124	1	US-08-449-045C-4
19	94	88.7	124	1	US-08-449-045C-4
20	94	88.7	124	6	5223610-3
21	51	48.1	448	4	US-08-878-989-2
22	51	48.1	448	4	US-09-272-796-2
23	51	48.1	508	4	US-09-344-700-4
24	42	39.6	835	2	US-09-027-337-2
25	41.5	39.2	1090	4	US-09-346-237-5
26	40.5	38.2	856	3	US-08-709-784-2
27	40.5	38.2	862	2	US-08-209-521-23

28	40.5	38.2	862	2	US-08-209-521-30	Sequence 30, Appl
29	40.5	38.2	862	4	US-09-059-461-2	Sequence 2, Appl
30	40.5	38.2	862	4	US-08-961-810-133	Sequence 133, App
31	40.5	38.2	862	4	US-08-352-902D-133	Sequence 133, App
32	39	36.8	251	1	US-07-956-700B-94	Sequence 94, Appl
33	39	36.8	251	1	US-08-476-537-94	Sequence 94, Appl
34	39	36.8	251	1	US-08-485-607-94	Sequence 94, Appl
35	39	36.8	251	2	US-08-475-879-94	Sequence 94, Appl
36	39	36.8	346	2	US-08-602-359A-34	Sequence 34, Appl
37	39	36.8	384	4	US-09-025-580-35	Sequence 35, Appl
38	39	36.8	427	4	US-09-025-580-36	Sequence 36, Appl
39	39	36.8	459	6	5194375-6	Patent No. 5194375
40	39	36.8	693	4	US-09-564-805-234	Sequence 234, App
41	39	36.8	775	2	US-08-714-070A-1	Sequence 1, Appl
42	38	35.8	845	1	US-08-416-950-11	Sequence 11, Appl
43	38	35.8	845	2	US-08-469-830-11	Sequence 11, Appl
44	38	35.8	1073	4	US-09-541-782-6	Sequence 6, Appl
45	38	35.8	1810	5	PCT-US95-11684-4	Sequence 4, Appl

ALIGNMENTS

```
RESULT 1
US-08-952-337-5
Sequence 5, Application US/08952337
Patent No. 6019973
GENERAL INFORMATION:
APPLICANT: Holmgren, Jan
TITLE OF INVENTION: HYBRID MOLECULES BETWEEN HEAT-LABILE
FILE REFERENCE: 3846/OD758
CURRENT APPLICATION NUMBER: US/08/952.337
EARLIER FILING DATE: 1998-01-05
EARLIER APPLICATION NUMBER: PCT/SE96/00570
EARLIER FILING DATE: 1996-05-02
EARLIER APPLICATION NUMBER: SE 9501682-0
EARLIER FILING DATE: 1995-05-05
NUMBER OF SEQ ID NOS: 6
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 5
LENGTH: 102
TYPE: PRT
ORGANISM: Vibrio cholerae
US-08-952-337-5

Query Match          100.0%; Score 106; DB 3; Length 102;
Best Local Similarity 100.0%; Pred. No. 2.5e-11;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      1 GATFOVEYVPGSQHDSOKKAI 21
Db      44 GATFOVEYVPGSQHDSOKKAI 64

RESULT 2
US-08-952-337-6
Sequence 6, Application US/08952337
Patent No. 6019973
GENERAL INFORMATION:
APPLICANT: Holmgren, Jan
TITLE OF INVENTION: HYBRID MOLECULES BETWEEN HEAT-LABILE
FILE REFERENCE: 3846/OD758
CURRENT APPLICATION NUMBER: US/08/952.337
EARLIER FILING DATE: 1998-01-05
EARLIER APPLICATION NUMBER: PCT/SE96/00570
EARLIER FILING DATE: 1996-05-02
EARLIER APPLICATION NUMBER: SE 9501682-0
EARLIER FILING DATE: 1995-05-05
```

NUMBER OF SEQ ID NOS: 6
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 6
LENGTH: 102
TYPE: PRT
ORGANISM: Escherichia coli
US-08-952-337-6

Query Match 100.0%; Score 106; DB 3; Length 102;
Best Local Similarity 100.0%; Pred. No. 2.5e-11;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GATFOVEVPGSQHIDSOKKAI 21
|||||
DB 44 GATFOVEVPGSQHIDSOKKAI 64

RESULT 3
US-08-472-171-2
Sequence 2, Application US/08472171
Patent No. 5932714
GENERAL INFORMATION:
APPLICANT: Loosmore, Sheena M.
APPLICANT: Yacoub, Reza K.
APPLICANT: Zealey, Gavin R.
APPLICANT: Klein, Michel H.
TITLE OF INVENTION: Expression Of Gene Products From
Genetically Manipulated Strains Of Bordetella
NUMBER OF SEQUENCES: 56
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sim & McBurney
STREET: 330 University Avenue, Suite 701
CITY: Toronto
STATE: Ontario
COUNTRY: Canada
ZIP: M5G 1R7
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/472,171
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/393,334
FILING DATE: 23-FEB-1995
ATTORNEY/AGENT INFORMATION:
NAME: Stewart, Michael I.
REGISTRATION NUMBER: 24,973
REFERENCE/DOCKET NUMBER: 1038-507 MIS:vg
TELECOMMUNICATION INFORMATION:
TELEPHONE: 416-595-1155
TELEFAX: 416-595-1163
TELEX: 065-24567 Simdas
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 103 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-472-171-2

Query Match 100.0%; Score 106; DB 2; Length 103;
Best Local Similarity 100.0%; Pred. No. 2.5e-11;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GATFOVEVPGSQHIDSOKKAI 21
|||||
DB 45 GATFOVEVPGSQHIDSOKKAI 65

RESULT 4
US-08-894-526-2
Sequence 2, Application US/08894526
Patent No. 5942418
GENERAL INFORMATION:
APPLICANT: Loosmore, Sheena M.
APPLICANT: Yacoub, Reza K.
APPLICANT: Zealey, Gavin R.
APPLICANT: Klein, Michel H.
TITLE OF INVENTION: EXPRESSION OF GENE PRODUCTS FROM
GENETICALLY MANIPULATED STRAINS OF BORDETELLA
NUMBER OF SEQUENCES: 57
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sim & McBurney
STREET: 6th Floor, 330 University Avenue
CITY: Toronto
STATE: Ontario
COUNTRY: Canada
ZIP: M5G 1R7
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/894,526
FILING DATE: 01-DEC-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Stewart, Michael I.
REGISTRATION NUMBER: 24,973
REFERENCE/DOCKET NUMBER: 1038-724 MIS:jb
TELECOMMUNICATION INFORMATION:
TELEPHONE: (416) 595-1155
TELEFAX: (416) 595-1163
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 103 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-894-526-2

Query Match 100.0%; Score 106; DB 2; Length 103;
Best Local Similarity 100.0%; Pred. No. 2.5e-11;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GATFOVEVPGSQHIDSOKKAI 21
|||||
DB 45 GATFOVEVPGSQHIDSOKKAI 65

RESULT 5
US-09-013-047-2
Sequence 2, Application US/09013047
Patent No. 5998168
GENERAL INFORMATION:
APPLICANT: Loosmore, Sheena M.
APPLICANT: Yacoub, Reza K.
APPLICANT: Zealey, Gavin R.
APPLICANT: Klein, Michel H.
TITLE OF INVENTION: Expression Of Gene Products From
Genetically Manipulated Strains Of Bordetella
NUMBER OF SEQUENCES: 56
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sim & McBurney
STREET: 330 University Avenue, 6th Floor
CITY: Toronto
STATE: Ontario
COUNTRY: Canada

ZIP: M5G 1R7
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/013,047
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/472,171
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/393,334
FILING DATE: 23-FEB-1995
ATTORNEY/AGENT INFORMATION:
NAME: Stewart, Michael I.
REGISTRATION NUMBER: 24,973
REFERENCE/DOCKET NUMBER: 1038-507 MIS:vg
TELECOMMUNICATION INFORMATION:
TELEPHONE: 416-595-1153
TELEFAX: 416-595-1153
TELEX: 065-24567 Simbas
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 103 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-013-047-2

Query Match 100.0%; Score 106; DB 2; Length 103;
Best Local Similarity 100.0%; Pred. No. 2.5e-11;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GATFOVEVPGSQHIDSOKKAI 21
|||||
DB 45 GATFOVEVPGSQHIDSOKKAI 65

RESULT 6
US-09-374-597-2
Sequence 2, Application US/09374597
Patent No. 6140082
GENERAL INFORMATION:
APPLICANT: Loosmore, Sheena M.
APPLICANT: Yacoob, Reza K.
APPLICANT: Zeale, Gavin R.
APPLICANT: Klein, Michael H.
TITLE OF INVENTION: Expression Of Gene Products From
NUMBER OF SEQUENCES: 56
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sim & McBurney
STREET: 330 University Avenue, 6th Floor
CITY: Toronto
STATE: Ontario
COUNTRY: Canada
ZIP: M5G 1R7
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/374,597
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/393,334
FILING DATE: FEBRUARY 23, 1995

CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Stewart, Michael I.
REGISTRATION NUMBER: 24973
REFERENCE/DOCKET NUMBER: 1038-964
TELECOMMUNICATION INFORMATION:
TELEPHONE: 416-595-1155
TELEFAX: 416-595-1163
TELEX: 065-24567 Simbas
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 103 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-374-597-2

Query Match 100.0%; Score 106; DB 4; Length 103;
Best Local Similarity 100.0%; Pred. No. 2.5e-11;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GATFOVEVPGSQHIDSOKKAI 21
|||||
DB 45 GATFOVEVPGSQHIDSOKKAI 65

RESULT 7
US-09-191-852-21
Sequence 21, Application US/09191852
Patent No. 6194560
GENERAL INFORMATION:
APPLICANT: Charles J. Arentzen, Hugh S. Mason, and Tariq A. Haq
TITLE OF INVENTION: ORAL IMMUNIZATION WITH TRANSGENIC PLANTS
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fulbright & Jaworski L.L.P.
STREET: 1301 McKinney, Suite 5100
CITY: Houston
STATE: TX
COUNTRY: USA
ZIP: 77010
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/191,852
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/13376
FILING DATE: 24-OCT-1995
APPLICATION NUMBER: 08/817,906
FILING DATE: 04-AUG-1997
ATTORNEY/AGENT INFORMATION:
NAME: Fox, David L.
REGISTRATION NUMBER: 40,612
REFERENCE/DOCKET NUMBER: P01590US1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 713-651-5151
TELEFAX: 713-651-5246
INFORMATION FOR SEQ ID NO: 21:
SEQUENCE CHARACTERISTICS:
LENGTH: 103 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-09-191-852-21

Query Match 100.0%; Score 106; DB 4; Length 103;
Best Local Similarity 100.0%; Pred. No. 2.5e-11;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GATFOVEVPGSOHIDSOKKAI 21
|||||

Db 45 GATFOVEVPGSOHIDSOKKAI 65

RESULT 8
PCT-US95-13376-21
; Sequence 21, Application PC/TUS9513376
; GENERAL INFORMATION:
; APPLICANT: The Texas A&M University System
; APPLICANT: 310 Wisenbaker
; APPLICANT: College Station, Texas 77843-3369
; TITLE OF INVENTION: ORAL IMMUNIZATION WITH TRANSGENIC PLANTS
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pravel, Hewitt, Kimball & Krieger
; STREET: 1177 West Loop South, 10th Floor
; CITY: Houston
; STATE: TX
; COUNTRY: USA
; ZIP: 77027-9095
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/13376
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/328,716
; FILING DATE: 24-OCT-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Jones, John W.
; REGISTRATION NUMBER: 31,380
; REFERENCE/DOCKET NUMBER: 36170/3P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 713-850-0909
; TELEFAX: 713-850-0165
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 103 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; PCT-US95-13376-21

Query Match 100.0%; Score 106; DB 5; Length 103;
Best Local Similarity 100.0%; Pred. No. 2.5e-11;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GATFOVEVPGSOHIDSOKKAI 21
|||||

Db 45 GATFOVEVPGSOHIDSOKKAI 65

RESULT 9
US-08-952-337-1
; Sequence 1, Application US/08952337
; Patent No. 601973
; GENERAL INFORMATION:
; APPLICANT: Holmgren, Jan
; APPLICANT: Lebens, Michael R.
; TITLE OF INVENTION: HYBRID MOLECULES BETWEEN HEAT-LABILE
; TITLE OF INVENTION: ENTEROTOXIN AND CHOLERA TOXIN B SUBUNITS
; FILE REFERENCE: 3846/0D758
; CURRENT APPLICATION NUMBER: US/08/952,337
; CURRENT FILING DATE: 1998-01-05
; EARLIER APPLICATION NUMBER: PCT/SE96/00570
; EARLIER FILING DATE: 1996-05-02

; EARLIER APPLICATION NUMBER: SE 9501682-0
; EARLIER FILING DATE: 1995-05-05
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 123
; TYPE: PRT
; ORGANISM: Vibrio cholerae
US-08-952-337-1

Query Match 100.0%; Score 106; DB 3; Length 123;
Best Local Similarity 100.0%; Pred. No. 3.1e-11;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GATFOVEVPGSOHIDSOKKAI 21
|||||

Db 65 GATFOVEVPGSOHIDSOKKAI 85

RESULT 10
US-08-952-337-2
; Sequence 2, Application US/08952337
; Patent No. 601973
; GENERAL INFORMATION:
; APPLICANT: Holmgren, Jan
; APPLICANT: Lebens, Michael R.
; TITLE OF INVENTION: HYBRID MOLECULES BETWEEN HEAT-LABILE
; TITLE OF INVENTION: ENTEROTOXIN AND CHOLERA TOXIN B SUBUNITS
; FILE REFERENCE: 3846/0D758
; CURRENT APPLICATION NUMBER: US/08/952,337
; CURRENT FILING DATE: 1998-01-05
; EARLIER APPLICATION NUMBER: PCT/SE96/00570
; EARLIER FILING DATE: 1996-05-02
; EARLIER APPLICATION NUMBER: SE 9501682-0
; EARLIER FILING DATE: 1995-05-05
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 123
; TYPE: PRT
; ORGANISM: Escherichia coli
US-08-952-337-2

Query Match 100.0%; Score 106; DB 3; Length 123;
Best Local Similarity 100.0%; Pred. No. 3.1e-11;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GATFOVEVPGSOHIDSOKKAI 21
|||||

Db 65 GATFOVEVPGSOHIDSOKKAI 85

RESULT 11
US-08-747-410-2
; Sequence 2, Application US/08747410
; Patent No. 5993820
; GENERAL INFORMATION:
; APPLICANT: BACDASARIAN, Michael
; APPLICANT: IRELAND, James
; TITLE OF INVENTION: CHIMERIC LTB VACCINES
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Merchant, Gould, Smith, Edell, Welter & Schmidt
; STREET: 3100 No. 5993820 West Center, 90 South Seventh St.
; CITY: Minneapolis
; STATE: MN
; COUNTRY: USA
; ZIP: 55402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible

```

Query Match          100.0%; Score 106; DB 2; Length 124;
Best Local Similarity 100.0%; Pred. No. 3 2e-11;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      1 GATFOVEVPSQHIDSOKKAI 21
      |||||||
Db      66 GATFOVEVPSQHIDSOKKAI 86

RESULT 12
US-08-292-968-26
: Sequence 26, Application US/08292968
: Patent No. 5856122
: GENERAL INFORMATION:
: APPLICANT: READ, Randy J.
: APPLICANT: STEIN, Penelope E.
: APPLICANT: COCKLE, Stephen A.
: APPLICANT: COMER, Raymond P.
: APPLICANT: LOOSMORE, Sheena
: APPLICANT: KLEIN, Michel H.
: APPLICANT: ARMSTRONG, Glen D.
: APPLICANT: HAZES, Bart
: TITLE OF INVENTION: MODIFICATION OF PERTUSSIS TOXIN
: NUMBER OF SEQUENCES: 46
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Slim & Mcburney
: STREET: Suite 701, 330 University Avenue
: CITY: Toronto
: STATE: Ontario
: COUNTRY: Canada
: ZIP: M5G 1R7
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentln Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/292,968
: FILING DATE:
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/110,947
:

```

```

Query Match          95.3%; Score 101; DB 2; Length 93;
Best Local Similarity 95.2%; Pred. No. 1.7e-10;
Matches      20; Conservative    0; Mismatches     1; Indels       0; Gaps        0;

OY              1   GATFOVEVPGSQHIDSOKKAI 21
                  | |||||
Db             35   GETFOVEVPGSQHIDSOKKAI 55

RESULT         13
US-08-467-974 -26
Sequence 26, Application US/08467974
Patent No. 3963385
GENERAL INFORMATION:
APPLICANT: READ, Randy J.
APPLICANT: STEIN, Penelope E.
APPLICANT: COCKLE, Stephen A.
APPLICANT: COMEN, Raymond P.
APPLICANT: LOOSMORE, Sheena
APPLICANT: KLEIN, Michel H.
APPLICANT: ARMSTRONG, Glen D.
TITLE OF INVENTION: MODIFICATION OF PERTUSSIS TOXIN
NUMBER OF SEQUENCES: 46
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sim & McBurney
STREET: Suite 701, 330 University Avenue
CITY: Toronto
STATE: Ontario
COUNTRY: Canada
ZIP: M5G 1R7

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/467,974
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/467,536
FILING DATE: 06-JUN-1995
APPLICATION NUMBER: US 08/292,968
FILING DATE: 22-AUG-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/251,121
FILING DATE: 31-MAY-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/110,947
FILING DATE: 24-AUG-1993
ATTORNEY/AGENT INFORMATION:
NAME: STEWART, Michael I.

```

REGISTRATION NUMBER: 24,973
REFERENCE/DOCKET NUMBER: 1038-454 MIS:vg
TELECOMMUNICATION INFORMATION:
TELEPHONE: (416) 595-1155
TELEFAX: (416) 595-1163
INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 93 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-467-974-26

Query Match 95.3%; Score 101; DB 2; Length 93;
Best Local Similarity 95.2%; Pred. No. 1.7e-10;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GATFQVEVPGSQHDSQKAI 21
Db 35 GETFQVEVPGSQHDSQKAI 55

RESULT 14
US-08-467-536-26
Sequence 26, Application US/08467536
Patent No. 5977304
GENERAL INFORMATION:
APPLICANT: READ, Randy J.
APPLICANT: STEIN, Penelope E.
APPLICANT: COCKLE, Stephen A.
APPLICANT: OOMEN, Raymond P.
APPLICANT: LOOSMORE, Sheena
APPLICANT: KLEIN, Michel H.
APPLICANT: ARMSTRONG, Glen D.
TITLE OF INVENTION: MODIFICATION OF PERTUSSIS TOXIN
NUMBER OF SEQUENCES: 46
CORRESPONDENCE ADDRESS:
ADDRESSEE: Slim & McBurney
STREET: Suite 701, 330 University Avenue
City: Toronto
STATE: Ontario
COUNTRY: Canada
ZIP: M5G 1R7
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/467,536
FILING DATE: 06-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/292,968
FILING DATE: 22-AUG-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/251,121
FILING DATE: 31-MAY-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/110,947
FILING DATE: 24-AUG-1993
ATTORNEY/AGENT INFORMATION:
NAME: STEWART, Michael I.
REGISTRATION NUMBER: 24,973
REFERENCE/DOCKET NUMBER: 1038-455 MIS:vg
TELECOMMUNICATION INFORMATION:
TELEPHONE: (416) 595-1155
TELEFAX: (416) 595-1163
INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 93 amino acids

TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-467-536-26

Query Match 95.3%; Score 101; DB 2; Length 93;
Best Local Similarity 95.2%; Pred. No. 1.7e-10;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GATFQVEVPGSQHDSQKAI 21
Db 35 GETFQVEVPGSQHDSQKAI 55

RESULT 15
US-08-467-976-26
Sequence 26, Application US/08467976
Patent No. 6018022
GENERAL INFORMATION:
APPLICANT: READ, Randy J.
APPLICANT: STEIN, Penelope E.
APPLICANT: COCKLE, Stephen A.
APPLICANT: OOMEN, Raymond P.
APPLICANT: LOOSMORE, Sheena
APPLICANT: KLEIN, Michel H.
APPLICANT: ARMSTRONG, Glen D.
TITLE OF INVENTION: MODIFICATION OF PERTUSSIS TOXIN
NUMBER OF SEQUENCES: 46
CORRESPONDENCE ADDRESS:
ADDRESSEE: Slim & McBurney
STREET: Suite 701, 330 University Avenue
City: Toronto
STATE: Ontario
COUNTRY: Canada
ZIP: M5G 1R7
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/467,976
FILING DATE: 06-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/292,968
FILING DATE: 22-AUG-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/251,121
FILING DATE: 31-MAY-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/110,947
FILING DATE: 24-AUG-1993
ATTORNEY/AGENT INFORMATION:
NAME: STEWART, Michael I.
REGISTRATION NUMBER: 24,973
REFERENCE/DOCKET NUMBER: 1038-453 MIS:vg
TELECOMMUNICATION INFORMATION:
TELEPHONE: (416) 595-1155
TELEFAX: (416) 595-1163
INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 93 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-467-976-26

Query Match 95.3%; Score 101; DB 3; Length 93;
Best Local Similarity 95.2%; Pred. No. 1.7e-10;

Mon Jul 8 07:51:53 2002

us-09-786-648-4.ra1

Page 7

Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 GATFOVEVPGSQHIDSOKKAI 21

Db 35 GETFOVEVPGSQHIDSOKKAI 55

Search completed: July 3, 2002, 08:55:39
Job time: 214 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 3, 2002, 09:03:20 ; Search time 401.04 Seconds
(without alignments)
18,431 Million cell updates/sec

Title: US-09-786-648-4
Perfect score: 106
Sequence: 1 GATFOVEYPSQSHDSOKKAI 21

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 3502263 seqs, 351980561 residues
Total number of hits satisfying chosen parameters: 3502263

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Pending Patents, AA Main:*

1: /cgn2_6/ptodata/2/paa/US06_COMB.pep.*
2: /cgn2_6/ptodata/2/paa/US07_COMB.pep.*
3: /cgn2_6/ptodata/2/paa/US07_COMB.pep.*
4: /cgn2_6/ptodata/2/paa/US08_COMB.pep.*
5: /cgn2_6/ptodata/2/paa/US08_COMB.pep.*
6: /cgn2_6/ptodata/2/paa/US08_COMB.pep.*
7: /cgn2_6/ptodata/2/paa/US08_COMB.pep.*
8: /cgn2_6/ptodata/2/paa/US08_COMB.pep.*
9: /cgn2_6/ptodata/2/paa/US08_COMB.pep.*
10: /cgn2_6/ptodata/2/paa/US08_COMB.pep.*
11: /cgn2_6/ptodata/2/paa/US08_COMB.pep.*
12: /cgn2_6/ptodata/2/paa/US08_COMB.pep.*
13: /cgn2_6/ptodata/2/paa/US08_COMB.pep.*
14: /cgn2_6/ptodata/2/paa/US09_COMB.pep.*
15: /cgn2_6/ptodata/2/paa/US09_COMB.pep.*
16: /cgn2_6/ptodata/2/paa/US09_COMB.pep.*
17: /cgn2_6/ptodata/2/paa/US09_COMB.pep.*
18: /cgn2_6/ptodata/2/paa/US09_COMB.pep.*
19: /cgn2_6/ptodata/2/paa/US09_COMB.pep.*
20: /cgn2_6/ptodata/2/paa/US09_COMB.pep.*
21: /cgn2_6/ptodata/2/paa/US09_COMB.pep.*
22: /cgn2_6/ptodata/2/paa/US09_COMB.pep.*
23: /cgn2_6/ptodata/2/paa/US09_COMB.pep.*
24: /cgn2_6/ptodata/2/paa/US10_COMB.pep.*
25: /cgn2_6/ptodata/2/paa/US10_COMB.pep.*
26: /cgn2_6/ptodata/2/paa/US60_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	106	100.0	21	US-09-786-648-4	Sequence 4, Appli
2	106	100.0	103	US-08-393-334-2	Sequence 2, Appli
3	106	100.0	103	US-08-782-832-15	Sequence 15, Appl
4	106	100.0	103	US-08-817-906-21	Sequence 21, Appl
5	106	100.0	103	US-09-836-433-14	Sequence 14, Appl
6	106	100.0	116	US-09-836-433-20	Sequence 20, Appl
7	106	100.0	119	US-09-836-433-22	Sequence 22, Appl

8	106	100.0	124	1	PCT-US99-30747-55	Sequence 55, Appl
9	106	100.0	124	18	US-09-470-124-55	Sequence 55, Appl
10	106	100.0	138	13	US-08-914-479-2	Sequence 2, Appli
11	106	100.0	138	13	US-08-914-479A-2	Sequence 2, Appli
12	106	100.0	313	21	US-09-756-983-15	Sequence 15, Appl
13	106	100.0	313	21	US-09-756-983-18	Sequence 18, Appl
14	106	100.0	364	21	US-09-756-983-22	Sequence 22, Appl
15	101	95.3	21	21	US-09-786-648-5	Sequence 21, Appl
16	101	95.3	93	5	US-08-110-947-10	Sequence 10, Appl
17	101	95.3	93	5	US-08-110-947A-26	Sequence 26, Appl
18	101	95.3	93	6	US-08-251-121-26	Sequence 26, Appl
19	101	95.3	371	5	US-08-150-305A-3	Sequence 3, Appli
20	101	95.3	371	11	US-08-784-218-6	Sequence 6, Appli
21	101	95.3	371	12	US-08-829-026-5	Sequence 5, Appli
22	100	94.3	124	1	PCT-US99-30747-57	Sequence 7, Appl
23	100	94.3	124	18	US-09-470-124-57	Sequence 57, Appl
24	94	88.7	124	21	US-09-760-234-7	Sequence 7, Appli
25	94	88.7	382	1	PCT-US01-08582-3	Sequence 3, Appli
26	94	88.7	382	1	PCT-US01-08582-4	Sequence 4, Appli
27	94	88.7	461	14	US-09-051-315-2	Sequence 2, Appli
28	94	88.7	461	18	US-09-423-493-2	Sequence 2, Appli
29	94	88.7	750	18	US-09-402-100-2	Sequence 2, Appli
30	94	88.7	1338	18	US-09-402-100-4	Sequence 4, Appli
31	76	71.7	15	11	US-08-732-371-1	Sequence 1, Appli
32	76	71.7	15	11	US-08-732-371A-1	Sequence 1, Appli
33	62	58.5	12	21	US-09-786-648-3	Sequence 3, Appli
34	51	48.1	266	21	US-09-758-445-415	Sequence 415, App
35	51	48.1	448	21	US-09-769-970-2	Sequence 2, Appli
36	51	48.1	508	19	US-09-563-997-4	Sequence 4, Appli
37	48	45.3	484	16	US-09-252-991A-29252	Sequence 29252, A
38	47	44.3	369	21	US-09-708-427-58248	Sequence 58248, A
39	47	44.3	395	26	US-60-324-109-19516	Sequence 19516, A
40	47	44.3	433	21	US-09-708-427-58247	Sequence 58247, A
41	47	44.3	494	21	US-09-708-427-58246	Sequence 58246, A
42	46	43.4	99	21	US-09-758-463-761	Sequence 761, App
43	46	43.4	101	15	US-09-134-000-3773	Sequence 3773, Ap
44	46	43.4	365	20	US-09-692-058-19	Sequence 19, Appl
45	46	43.4	529	18	US-09-417-507-37800	Sequence 37800, A

ALIGNMENTS

RESULT 1

US-09-786-648-4

Sequence 4, Application US/09786648

GENERAL INFORMATION:

APPLICANT: Williams, Neil Andrew

TITLE OF INVENTION: Peptide Fragments of Cholera Toxin B or Enterotoxin B as Vacci

FILE REFERENCE: 7438

ADJUVANTS

CURRENT APPLICATION NUMBER: US/09/786,648

CURRENT FILING DATE: 2001-03-07

PRIOR APPLICATION NUMBER: PCT/GB99/02970

PRIOR FILING DATE: 1999-09-07

NUMBER OF SEQ ID NOS: 6

SOFTWARE: MS DOS

SEQ ID NO 4

LENGTH: 21

TYPE: PRT

ORGANISM: E. coli

FEATURE:

LOCATION: 45...65

OTHER INFORMATION: isolated or synthetic E. coli

OTHER INFORMATION: human variant E. coli

US-09-786-648-4

Query Match 100.0%; Score 106; DB 21; Length 21;
Best Local Similarity 100.0%; Pred. No. 1.3e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GATFOVEVPGSOHIDSOKKAI 21
|||||
DB 1 GATFOVEVPGSOHIDSOKKAI 21

RESULT 2

US-08-393-334-2
Sequence 2, Application US/08393334
GENERAL INFORMATION:
APPLICANT: Loosmore, Sheena M.
APPLICANT: Yacoub, Reza K.
APPLICANT: Zealey, Gavin R.
APPLICANT: Klein, Michel H.
TITLE OF INVENTION: Expression Of Gene Products From
NUMBER OF SEQUENCES: 56
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Slim & Mcburey
STREET: 330 University Avenue, Suite 701
CITY: Toronto
STATE: Ontario
COUNTRY: Canada
ZIP: M5G 1R7
COMPUTER READABLE FORM:
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/393,334
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Stewart, Michael I
REGISTRATION NUMBER: 24973
REFERENCE/DOCKET NUMBER: 1038-417
TELECOMMUNICATION INFORMATION:
TELEPHONE: 416-595-1155
TELEFAX: 416-595-1163
TELEX: 065-24567 Simbas
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 103 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-393-334-2

Query Match 100.0%; Score 106; DB 7; Length 103;
Best Local Similarity 100.0%; Pred. No. 9.8e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GATFOVEVPGSOHIDSOKKAI 21
|||||
DB 45 GATFOVEVPGSOHIDSOKKAI 65

RESULT 3

US-08-782-832-15
Sequence 15, Application US/08782832
GENERAL INFORMATION:
APPLICANT: Arntzen, Charles J.
APPLICANT: Mason, Hugh S.
APPLICANT: Haq, Tariq A.
TITLE OF INVENTION: PRODUCTION OF AN ORALLY IMMUNOGENIC
BACTERIAL PROTEIN IN TRANSGENIC PLANTS
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Pravel, Hewitt, Kimball & Krieger
STREET: 1177 West Loop South, 10th floor
CITY: Houston
STATE: TX

COUNTRY: USA
ZIP: 77027-9095
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/782,832
FILING DATE: 13-JAN-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/328,716
FILING DATE: 24-OCT-1994
ATTORNEY/AGENT INFORMATION:
NAME: Kettelberger, Denise M.
REGISTRATION NUMBER: 33,924
REFERENCE/DOCKET NUMBER: 36170/3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 713-850-0909
TELEFAX: 713-850-0165
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 103 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-782-832-15

Query Match 100.0%; Score 106; DB 11; Length 103;
Best Local Similarity 100.0%; Pred. No. 9.8e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GATFOVEVPGSOHIDSOKKAI 21
|||||
DB 45 GATFOVEVPGSOHIDSOKKAI 65

RESULT 4

US-08-817-906-21
Sequence 21, Application US/08817906
GENERAL INFORMATION:
APPLICANT: Charles J. Arntzen, Hugh S. Mason, John D. Clements,
APPLICANT: and Tariq A. Haq
TITLE OF INVENTION: ORAL IMMUNIZATION WITH TRANSGENIC PLANTS
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Fulbright & Jaworski L.L.P.
STREET: 1301 McKinney, Suite 5100
CITY: Houston
STATE: TX
COUNTRY: USA
ZIP: 77010
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/817,906
FILING DATE: 08/04/97
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/13376
FILING DATE: 24-OCT-1995
ATTORNEY/AGENT INFORMATION:
NAME: Fox, David L.
REGISTRATION NUMBER: 40,612
REFERENCE/DOCKET NUMBER: 36170/3P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 713-651-5151
TELEFAX: 713-651-5246
INFORMATION FOR SEQ ID NO: 21:

SEQUENCE CHARACTERISTICS:
 ; LENGTH: 103 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 US-08-817-906-21

Query Match 100.0%; Score 106; DB 12; Length 103;
 Best Local Similarity 100.0%; Pred. No. 9.8e-10;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GATFOVEVPGSOHIDSOKKAI 21
 ||||||||||||||||
 DB 45 GATFOVEVPGSOHIDSOKKAI 65

RESULT 5
 US-09-836-433-14

; Sequence 14, Application US/09836433
 ; GENERAL INFORMATION:
 ; APPLICANT: Yuki, Yoshikazu
 ; TITLE OF INVENTION: HYBRID PROTEINS FOR AUTOIMMUNE DISEASE
 ; FILE REFERENCE: Not Assigned
 ; CURRENT APPLICATION NUMBER: US/09/836,433
 ; CURRENT FILING DATE: 2001-04-16
 ; NUMBER OF SEQ ID NOS: 49
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 14
 ; LENGTH: 103
 ; TYPE: PRT
 ; ORGANISM: Vibrio cholerae
 US-09-836-433-14

Query Match 100.0%; Score 106; DB 22; Length 103;
 Best Local Similarity 100.0%; Pred. No. 9.8e-10;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GATFOVEVPGSOHIDSOKKAI 21
 ||||||||||||||||
 DB 45 GATFOVEVPGSOHIDSOKKAI 65

RESULT 6
 US-09-836-433-20

; Sequence 20, Application US/09836433
 ; GENERAL INFORMATION:
 ; APPLICANT: Yuki, Yoshikazu
 ; TITLE OF INVENTION: HYBRID PROTEINS FOR AUTOIMMUNE DISEASE
 ; FILE REFERENCE: Not Assigned
 ; CURRENT APPLICATION NUMBER: US/09/836,433
 ; CURRENT FILING DATE: 2001-04-16
 ; NUMBER OF SEQ ID NOS: 49
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 20
 ; LENGTH: 116
 ; TYPE: PRT
 ; ORGANISM: synthetic construct
 US-09-836-433-20

Query Match 100.0%; Score 106; DB 22; Length 116;
 Best Local Similarity 100.0%; Pred. No. 1.1e-09;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GATFOVEVPGSOHIDSOKKAI 21
 ||||||||||||||||
 DB 45 GATFOVEVPGSOHIDSOKKAI 65

RESULT 7

US-09-836-433-22

; Sequence 22, Application US/09836433
 ; GENERAL INFORMATION:
 ; APPLICANT: Yuki, Yoshikazu
 ; TITLE OF INVENTION: HYBRID PROTEINS FOR AUTOIMMUNE DISEASE
 ; FILE REFERENCE: Not Assigned
 ; CURRENT APPLICATION NUMBER: US/09/836,433
 ; CURRENT FILING DATE: 2001-04-16
 ; NUMBER OF SEQ ID NOS: 49
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 22
 ; LENGTH: 119
 ; TYPE: PRT
 ; ORGANISM: synthetic construct
 US-09-836-433-22

Query Match 100.0%; Score 106; DB 22; Length 119;
 Best Local Similarity 100.0%; Pred. No. 1.2e-09;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GATFOVEVPGSOHIDSOKKAI 21
 ||||||||||||||||
 DB 45 GATFOVEVPGSOHIDSOKKAI 65

RESULT 8
 PCT-US99-30747-55

; Sequence 55, Application PCT/US9930747
 ; GENERAL INFORMATION:
 ; APPLICANT: Boyce Thompson Institute for Plant Research at Cor
 ; TITLE OF INVENTION: Orally Immunogenic Bacterial Enterotoxins Expressed In
 ; FILE REFERENCE: 4868/85427
 ; CURRENT APPLICATION NUMBER: PCT/US99/30747
 ; CURRENT FILING DATE: 1999-12-22
 ; NUMBER OF SEQ ID NOS: 67
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 55
 ; LENGTH: 124
 ; TYPE: PRT
 ; ORGANISM: Escherichia coli
 PCT-US99-30747-55

Query Match 100.0%; Score 106; DB 1; Length 124;
 Best Local Similarity 100.0%; Pred. No. 1.2e-09;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GATFOVEVPGSOHIDSOKKAI 21
 ||||||||||||||||
 DB 66 GATFOVEVPGSOHIDSOKKAI 86

RESULT 9
 US-09-470-124-55

; Sequence 55, Application US/09470124
 ; GENERAL INFORMATION:
 ; APPLICANT: Arntzen
 ; TITLE OF INVENTION: Orally Immunogenic Bacterial Enterotoxins Expressed In
 ; FILE REFERENCE: 4868/84454
 ; CURRENT APPLICATION NUMBER: US/09/470,124
 ; CURRENT FILING DATE: 1999-12-24
 ; NUMBER OF SEQ ID NOS: 67
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 55
 ; LENGTH: 124
 ; TYPE: PRT
 ; ORGANISM: Escherichia coli
 US-09-470-124-55

Query Match 100.0%; Score 106; DB 18; Length 124;
Best Local Similarity 100.0%; Pred. No. 1.2e-09;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GATFOVEVPGSQHIDSOKKAI 21
|||||
Db 66 GATFOVEVPGSQHIDSOKKAI 86

RESULT 10
US-08-914-479-2
; Sequence 2, Application US/08914479
; GENERAL INFORMATION:
; APPLICANT: Date, James B.
; TITLE OF INVENTION: ANTIGEN OF HYBRID M PROTEIN AND CARRIER
; TITLE OF INVENTION: FOR GROUP A STREPTOCOCCAL VACCINE
; FILE REFERENCE: 48112.404C2
; CURRENT APPLICATION NUMBER: US/08/914,479
; PRIOR FILING DATE: 1997-08-19
; PRIOR APPLICATION NUMBER: 08/409,270
; PRIOR FILING DATE: 1995-03-23
; PRIOR APPLICATION NUMBER: 07/945,860
; PRIOR FILING DATE: 1992-09-16
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 138
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: LT-B-M24 hybrid molecule
US-08-914-479-2

Query Match 100.0%; Score 106; DB 13; Length 138;
Best Local Similarity 100.0%; Pred. No. 1.4e-09;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GATFOVEVPGSQHIDSOKKAI 21
|||||
Db 66 GATFOVEVPGSQHIDSOKKAI 86

RESULT 11
US-08-914-479A-2
; Sequence 2, Application US/08914479A
; GENERAL INFORMATION:
; APPLICANT: Date, James B.
; TITLE OF INVENTION: ANTIGEN OF HYBRID M PROTEIN AND CARRIER
; TITLE OF INVENTION: FOR GROUP A STREPTOCOCCAL VACCINE
; FILE REFERENCE: 48112.404C2
; CURRENT APPLICATION NUMBER: US/08/914,479A
; PRIOR FILING DATE: 1997-08-19
; PRIOR APPLICATION NUMBER: 08/409,270
; PRIOR FILING DATE: 1995-03-23
; PRIOR APPLICATION NUMBER: 07/945,860
; PRIOR FILING DATE: 1992-09-16
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 138
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: LT-B-M24 hybrid molecule
US-08-914-479A-2

Query Match 100.0%; Score 106; DB 13; Length 138;
Best Local Similarity 100.0%; Pred. No. 1.4e-09;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GATFOVEVPGSQHIDSOKKAI 21
|||||
Db 66 GATFOVEVPGSQHIDSOKKAI 86

RESULT 12
US-09-756-983-15
; Sequence 15, Application US/09756983
; GENERAL INFORMATION:
; APPLICANT: Albani, Salvatore
; TITLE OF INVENTION: METHODS FOR ISOLATION, QUANTIFICATION,
; TITLE OF INVENTION: CHARACTERIZATION AND MODULATION OF
; FILE REFERENCE: 246/285-CIP
; CURRENT APPLICATION NUMBER: US/09/756,983
; PRIOR FILING DATE: 2001-01-09
; PRIOR APPLICATION NUMBER: 60/105,018
; PRIOR FILING DATE: 1998-10-20
; PRIOR APPLICATION NUMBER: 09/421,506
; PRIOR FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: PCT/US99/2466
; PRIOR FILING DATE: 1999-10-19
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15
; LENGTH: 313
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Fusion constructs with human and bacterial sequences
US-09-756-983-15

Query Match 100.0%; Score 106; DB 21; Length 313;
Best Local Similarity 100.0%; Pred. No. 4e-09;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GATFOVEVPGSQHIDSOKKAI 21
|||||
Db 255 GATFOVEVPGSQHIDSOKKAI 275

RESULT 13
US-09-756-983-18
; Sequence 18, Application US/09756983
; GENERAL INFORMATION:
; APPLICANT: Albani, Salvatore
; TITLE OF INVENTION: METHODS FOR ISOLATION, QUANTIFICATION,
; TITLE OF INVENTION: CHARACTERIZATION AND MODULATION OF
; FILE REFERENCE: 246/285-CIP
; CURRENT APPLICATION NUMBER: US/09/756,983
; PRIOR FILING DATE: 2001-01-09
; PRIOR APPLICATION NUMBER: 60/105,018
; PRIOR FILING DATE: 1998-10-20
; PRIOR APPLICATION NUMBER: 09/421,506
; PRIOR FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: PCT/US99/2466
; PRIOR FILING DATE: 1999-10-19
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 18
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Fusion constructs with human and bacterial sequences
US-09-756-983-18

Query Match 100.0%; Score 106; DB 21; Length 351;
Best Local Similarity 100.0%; Pred. No. 4.6e-09;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GATFOVEVPGSQHIDSOKKAI 21
Db 293 GATFOVEVPGSQHIDSOKKAI 313

RESULT 14
US-09-756-983-22

; Sequence 22, Application US/09756983
; GENERAL INFORMATION:
; APPLICANT: Aldani, Salvatore
; TITLE OF INVENTION: METHODS FOR ISOLATION, QUANTIFICATION,
; TITLE OF INVENTION: CHARACTERIZATION AND MODULATION OF
; TITLE OF INVENTION: ANTIGEN-SPECIFIC T CELLS
; FILE REFERENCE: 246/285-CIP
; CURRENT APPLICATION NUMBER: US/09/756,983
; CURRENT FILING DATE: 2001-01-09
; PRIOR APPLICATION NUMBER: 60/105,018
; PRIOR FILING DATE: 1998-10-20
; PRIOR APPLICATION NUMBER: 09/421,506
; PRIOR FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: PCT/US99/2466
; PRIOR FILING DATE: 1999-10-19
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 22
; LENGTH: 364
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Fusion constructs with human and bacterial sequences
US-09-756-983-22

Query Match 100.0%; Score 106; DB 21; Length 364;
Best Local Similarity 100.0%; Pred. No. 4.9e-09;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GATFOVEVPGSQHIDSOKKAI 21
Db 306 GATFOVEVPGSQHIDSOKKAI 326

RESULT 15
US-09-786-648-5

; Sequence 5, Application US/09786648
; GENERAL INFORMATION:
; APPLICANT: Williams, Neil Andrew
; TITLE OF INVENTION: Peptide Fragments of Cholera Toxin B or Enterotoxin B as Vaccine
; FILE REFERENCE: 7438
; CURRENT APPLICATION NUMBER: US/09/786,648
; CURRENT FILING DATE: 2001-03-07
; PRIOR APPLICATION NUMBER: PCT/GB99/02970
; PRIOR FILING DATE: 1999-09-07
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: MS DOS
; SEQ ID NO 5
; LENGTH: 21
; TYPE: PRT
; ORGANISM: E. coli
; FEATURE:
; LOCATION: 45...65
; OTHER INFORMATION: isolated or synthetic EtXB beta4-alpha2 loop fragment derivable f
; OTHER INFORMATION: porcine E. coli
US-09-786-648-5

Query Match 95.3%; Score 101; DB 21; Length 21;
Best Local Similarity 95.2%; Pred. No. 9.3e-10;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GATFOVEVPGSQHIDSOKKAI 21
Db 1 GATFOVEVPGSQHIDSOKKAI 21

Search completed: July 3, 2002, 09:03:21
Job time: 621 sec

Mon Jul 8 07:51:54 2002

us-09-786-648-4.rapm

Page 6

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 3, 2002, 09:04:12 ; Search time 46.52 Seconds

(without alignments)
45.673 Million cell updates/sec

Title: US-09-786-648-4

Sequence: 1 GATFVEVPGSQHDSOKKAI 21

Scoring table:

BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 308740 seqs, 101176262 residues

Total number of hits satisfying chosen parameters: 308740

Minimum DB seq length: 0
Maximum DB seq length: 200000000Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Pending_Patents_AA.New:*
1: /cgn2_6/prodata/2/paa/PCT_NEW_COMB.pep:*
2: /cgn2_6/prodata/2/paa/US06_NEW_COMB.pep:*
3: /cgn2_6/prodata/2/paa/US07_NEW_COMB.pep:*
4: /cgn2_6/prodata/2/paa/US08_NEW_COMB.pep:*
5: /cgn2_6/prodata/2/paa/US09_NEW_COMB.pep:*
6: /cgn2_6/prodata/2/paa/US10_NEW_COMB.pep:*
7: /cgn2_6/prodata/2/paa/US60_NEW_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	106	100.0	103	6	US-10-110-364-10 Sequence 10, Appl
2	106	100.0	103	6	US-10-110-364-13 Sequence 13, Appl
3	106	100.0	105	6	US-10-110-364-12 Sequence 12, Appl
4	106	100.0	123	6	US-10-110-364-17 Sequence 17, Appl
5	106	100.0	124	6	US-10-110-364-2 Sequence 2, Appl
6	106	100.0	124	6	US-10-110-364-6 Sequence 6, Appl
7	106	100.0	124	6	US-10-110-364-11 Sequence 11, Appl
8	106	100.0	124	6	US-10-110-364-19 Sequence 19, Appl
9	106	100.0	124	6	US-10-110-364-21 Sequence 21, Appl
10	106	100.0	125	6	US-10-110-364-15 Sequence 15, Appl
11	106	100.0	138	6	US-10-141-627-2 Sequence 2, Appl
12	101	95.3	103	6	US-10-110-364-22 Sequence 22, Appl
13	101	95.3	123	6	US-10-110-364-16 Sequence 16, Appl
14	101	95.3	124	6	US-10-110-364-15 Sequence 15, Appl
15	101	95.3	124	6	US-10-110-364-18 Sequence 18, Appl
16	101	95.3	124	6	US-10-110-364-23 Sequence 23, Appl
17	100	94.3	103	6	US-10-110-364-8 Sequence 8, Appl
18	97	91.5	103	6	US-10-110-364-7 Sequence 7, Appl
19	94	88.7	124	6	US-10-110-364-5 Sequence 5, Appl
20	94	88.7	124	6	US-10-110-364-4 Sequence 4, Appl
21	94	88.7	124	6	US-10-110-364-9 Sequence 9, Appl
22	94	88.7	362	5	US-09-809-033A-3 Sequence 3, Appl
23	94	88.7	362	5	US-09-809-033A-4 Sequence 4, Appl
24	48	45.3	747	7	US-60-360-039-7808 Sequence 7808, Ap
25	43	40.6	84	5	US-09-620-393B-3233 Sequence 333, Ap
26	43	40.6	138	5	US-09-620-393B-3232 Sequence 3232, Ap

27	43	40.6	385	5	US-09-540-209B-5629 Sequence 5629, Ap
28	43	40.6	483	7	US-60-360-039-18750 Sequence 18750, A
29	42.5	40.1	376	7	US-60-360-039-21922 Sequence 21922, A
30	42	39.6	492	6	US-10-155-881-18930 Sequence 18930, A
31	42	39.6	782	1	PCT-US02-07826-312 Sequence 312, App
32	42	39.6	782	6	US-10-097-340-312 Sequence 312, App
33	42	39.6	855	1	PCT-US02-09039-2 Sequence 2, Appl
34	42	39.6	855	5	US-09-654-600A-2 Sequence 2, Appl
35	42	39.6	855	6	US-10-104-271-2 Sequence 2, Appl
36	42	39.6	855	6	US-10-099-700A-2 Sequence 2, Appl
37	42	39.6	890	7	US-60-360-039-3507 Sequence 3507, Ap
38	42	39.6	2359	6	US-10-155-881-28520 Sequence 28520, A
39	41	38.7	496	6	US-10-155-881-37499 Sequence 37499, A
40	41	38.7	549	5	US-09-540-209B-6846 Sequence 6846, Ap
41	40	37.7	192	7	US-60-360-039-10029 Sequence 10029, A
42	40	37.7	211	5	US-09-540-209B-9908 Sequence 9908, Ap
43	40	37.7	369	6	US-10-138-701-36 Sequence 36, Appl
44	40	37.7	417	5	US-09-540-209B-9466 Sequence 9466, Ap
45	40	37.7	657	7	US-60-360-039-5448 Sequence 5448, Ap

ALIGNMENTS

```
RESULT 1
US-10-110-364-10
; Sequence 10, Application US/10110364
; GENERAL INFORMATION:
; APPLICANT: Handley, Harold H.
; APPLICANT: Haaparanta, Tapio
; TITLE OF INVENTION: ABS TOXIN B SUBUNIT MUTANTS WITH ALTERED
; FILE REFERENCE: ACTBIO.004A
; CURRENT APPLICATION NUMBER: US/10/110.364
; CURRENT FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: PCT/US00/27607
; PRIOR FILING DATE: 2000-10-05
; PRIOR APPLICATION NUMBER: 60/158,561
; PRIOR FILING DATE: 1999-10-08
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 103
; TYPE: PRT
; ORGANISM: Vibrio cholera
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)..(103)
; OTHER INFORMATION: CTB variant from NCBI gene bank GI: 1421511
US-10-110-364-10

Query Match 100.0%; Score 106; DB 6; Length 103;
Best Local Similarity 100.0%; Pred. No. 5.2e-11;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GATFVEVPGSQHDSOKKAI 21
|||||
Db 45 GATFVEVPGSQHDSOKKAI 65

RESULT 2
US-10-110-364-13
; Sequence 13, Application US/10110364
; GENERAL INFORMATION:
; APPLICANT: Handley, Harold H.
; APPLICANT: Haaparanta, Tapio
; TITLE OF INVENTION: ABS TOXIN B SUBUNIT MUTANTS WITH ALTERED
; FILE REFERENCE: ACTBIO.004A
; CURRENT APPLICATION NUMBER: US/10/110.364
```

```
; CURRENT FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: PCT/US00/27607
; PRIOR FILING DATE: 2000-10-05
; PRIOR APPLICATION NUMBER: 60/158,561
; PRIOR FILING DATE: 1999-10-08
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 103
; TYPE: PRF
; ORGANISM: Vibrio cholera
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(103)
; OTHER INFORMATION: CTB variant from NCBI gene bank GI: 14215235
; OTHER INFORMATION: (Ogawa 41 R35D).
US-10-110-364-13
```

```
Query Match          100.0%; Score 106; DB 6; Length 103;
Best Local Similarity 100.0%; Pred. No. 5.2e-11;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 GATFOVEVPGSQHIDSOKKAI 21
DB 45 GATFOVEVPGSQHIDSOKKAI 65
```

```
RESULT 3
US-10-110-364-12
; Sequence 12, Application US/10110364
; GENERAL INFORMATION:
; APPLICANT: Handley, Harold H.
; APPLICANT: Haaparanta, Tapio
; APPLICANT: Ewalt, Karla L.
; TITLE OF INVENTION: ABS TOXIN B SUBUNIT MUTANTS WITH ALTERED
; TITLE OF INVENTION: CHEMICAL CONJUGATION CHARACTERISTICS
; FILE REFERENCE: ACTBIO.004A
; CURRENT APPLICATION NUMBER: US/10/110,364
; CURRENT FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: PCT/US00/27607
; PRIOR FILING DATE: 2000-10-05
; PRIOR APPLICATION NUMBER: 60/158,561
; PRIOR FILING DATE: 1999-10-08
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 105
; TYPE: PRF
; ORGANISM: Vibrio cholera
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(105)
; OTHER INFORMATION: CTB variant from NCBI gene bank GI: 2781121 (Ogawa
; OTHER INFORMATION: 41).
US-10-110-364-12
```

```
Query Match          100.0%; Score 106; DB 6; Length 105;
Best Local Similarity 100.0%; Pred. No. 5.3e-11;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 GATFOVEVPGSQHIDSOKKAI 21
DB 46 GATFOVEVPGSQHIDSOKKAI 66
```

```
RESULT 4
US-10-110-364-17
; Sequence 17, Application US/10110364
; GENERAL INFORMATION:
; APPLICANT: Handley, Harold H.
; APPLICANT: Haaparanta, Tapio
```

```
; APPLICANT: Ewalt, Karla L.
; TITLE OF INVENTION: ABS TOXIN B SUBUNIT MUTANTS WITH ALTERED
; TITLE OF INVENTION: CHEMICAL CONJUGATION CHARACTERISTICS
; FILE REFERENCE: ACTBIO.004A
; CURRENT APPLICATION NUMBER: US/10/110,364
; CURRENT FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: PCT/US00/27607
; PRIOR FILING DATE: 2000-10-05
; PRIOR APPLICATION NUMBER: 60/158,561
; PRIOR FILING DATE: 1999-10-08
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 123
; TYPE: PRF
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(123)
; OTHER INFORMATION: LTB variant from NCBI gene bank GI: 1395122.
US-10-110-364-17
```

```
Query Match          100.0%; Score 106; DB 6; Length 123;
Best Local Similarity 100.0%; Pred. No. 6.5e-11;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 GATFOVEVPGSQHIDSOKKAI 21
DB 66 GATFOVEVPGSQHIDSOKKAI 86
```

```
RESULT 5
US-10-110-364-2
; Sequence 2, Application US/10110364
; GENERAL INFORMATION:
; APPLICANT: Handley, Harold H.
; APPLICANT: Haaparanta, Tapio
; APPLICANT: Ewalt, Karla L.
; TITLE OF INVENTION: ABS TOXIN B SUBUNIT MUTANTS WITH ALTERED
; TITLE OF INVENTION: CHEMICAL CONJUGATION CHARACTERISTICS
; FILE REFERENCE: ACTBIO.004A
; CURRENT APPLICATION NUMBER: US/10/110,364
; CURRENT FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: PCT/US00/27607
; PRIOR FILING DATE: 2000-10-05
; PRIOR APPLICATION NUMBER: 60/158,561
; PRIOR FILING DATE: 1999-10-08
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 124
; TYPE: PRF
; ORGANISM: Vibrio cholera
US-10-110-364-2
```

```
Query Match          100.0%; Score 106; DB 6; Length 124;
Best Local Similarity 100.0%; Pred. No. 6.5e-11;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 GATFOVEVPGSQHIDSOKKAI 21
DB 66 GATFOVEVPGSQHIDSOKKAI 86
```

```
RESULT 6
US-10-110-364-6
; Sequence 6, Application US/10110364
; GENERAL INFORMATION:
; APPLICANT: Handley, Harold H.
; APPLICANT: Haaparanta, Tapio
; APPLICANT: Ewalt, Karla L.
```



```

; TITLE OF INVENTION: AB5 TOXIN B SUBUNIT MUTANTS WITH ALTERED
; TITLE OF INVENTION: CHEMICAL CONUGATION CHARACTERISTICS
; FILE REFERENCE: ACTBIO.004A
; CURRENT APPLICATION NUMBER: US/10/110,364
; CURRENT FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: PCT/US00/27607
; PRIOR FILING DATE: 2000-10-05
; PRIOR APPLICATION NUMBER: 60/158,561
; PRIOR FILING DATE: 1999-10-08
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 124
; TYPE: PRT
; ORGANISM: Vibrio cholera
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(124)
; OTHER INFORMATION: CTB variant from NCBI gene bank GI: 808900.
US-10-110-364-6
```

```

Query Match          100.0%; Score 106; DB 6; Length 124;
Best Local Similarity 100.0%; Pred. No. 6,5e-11;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 GATFOVEVPGSQHIDSOKKAI 21
Db      66 GATFOVEVPGSQHIDSOKKAI 86
```

```

RESULT 7
US-10-110-364-11
; Sequence 11, Application US/10110364
; GENERAL INFORMATION:
; APPLICANT: Handley, Harold H.
; APPLICANT: Haaparanta, Tapio
; APPLICANT: Ewalt, Karla L.
; TITLE OF INVENTION: AB5 TOXIN B SUBUNIT MUTANTS WITH ALTERED
; TITLE OF INVENTION: CHEMICAL CONUGATION CHARACTERISTICS
; FILE REFERENCE: ACTBIO.004A
; CURRENT APPLICATION NUMBER: US/10/110,364
; CURRENT FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: PCT/US00/27607
; PRIOR FILING DATE: 2000-10-05
; PRIOR APPLICATION NUMBER: 60/158,561
; PRIOR FILING DATE: 1999-10-08
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 124
; TYPE: PRT
; ORGANISM: Vibrio cholera
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(124)
; OTHER INFORMATION: CTB variant from NCBI gene bank GI: 48890 (classic)
; OTHER INFORMATION: 569B).
US-10-110-364-11
```

```

Query Match          100.0%; Score 106; DB 6; Length 124;
Best Local Similarity 100.0%; Pred. No. 6,5e-11;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 GATFOVEVPGSQHIDSOKKAI 21
Db      66 GATFOVEVPGSQHIDSOKKAI 86
```

```

RESULT 8
US-10-110-364-19
; Sequence 19, Application US/10110364
```

```

; GENERAL INFORMATION:
; APPLICANT: Handley, Harold H.
; APPLICANT: Haaparanta, Tapio
; APPLICANT: Ewalt, Karla L.
; TITLE OF INVENTION: AB5 TOXIN B SUBUNIT MUTANTS WITH ALTERED
; TITLE OF INVENTION: CHEMICAL CONUGATION CHARACTERISTICS
; FILE REFERENCE: ACTBIO.004A
; CURRENT APPLICATION NUMBER: US/10/110,364
; CURRENT FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: PCT/US00/27607
; PRIOR FILING DATE: 2000-10-05
; PRIOR APPLICATION NUMBER: 60/158,561
; PRIOR FILING DATE: 1999-10-08
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 19
; LENGTH: 124
; TYPE: PRT
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(124)
; OTHER INFORMATION: LTB variant from NCBI gene bank GI: 1648865
; OTHER INFORMATION: (LT87).
US-10-110-364-19
```

```

Query Match          100.0%; Score 106; DB 6; Length 124;
Best Local Similarity 100.0%; Pred. No. 6,5e-11;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 GATFOVEVPGSQHIDSOKKAI 21
Db      66 GATFOVEVPGSQHIDSOKKAI 86
```

```

RESULT 9
US-10-110-364-21
; Sequence 21, Application US/10110364
; GENERAL INFORMATION:
; APPLICANT: Handley, Harold H.
; APPLICANT: Haaparanta, Tapio
; APPLICANT: Ewalt, Karla L.
; TITLE OF INVENTION: AB5 TOXIN B SUBUNIT MUTANTS WITH ALTERED
; TITLE OF INVENTION: CHEMICAL CONUGATION CHARACTERISTICS
; FILE REFERENCE: ACTBIO.004A
; CURRENT APPLICATION NUMBER: US/10/110,364
; CURRENT FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: PCT/US00/27607
; PRIOR FILING DATE: 2000-10-05
; PRIOR APPLICATION NUMBER: 60/158,561
; PRIOR FILING DATE: 1999-10-08
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 21
; LENGTH: 124
; TYPE: PRT
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(124)
; OTHER INFORMATION: LTB variant from NCBI gene bank GI: 408996.
US-10-110-364-21
```

```

Query Match          100.0%; Score 106; DB 6; Length 124;
Best Local Similarity 100.0%; Pred. No. 6,5e-11;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 GATFOVEVPGSQHIDSOKKAI 21
Db      66 GATFOVEVPGSQHIDSOKKAI 86
```

RESULT 10
US-10-110-364-15
; Sequence 15, Application US/10110364
; GENERAL INFORMATION:
; APPLICANT: Handley, Harold H.
; APPLICANT: Haaparanta, Tapio
; APPLICANT: Ewalt, Karla L.
; TITLE OF INVENTION: ABS TOXIN B SUBUNIT MUTANTS WITH ALTERED
; FILE REFERENCE: ACTBIO.004A
; CURRENT APPLICATION NUMBER: US/10/110,364
; CURRENT FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: PCT/US00/27607
; PRIOR FILING DATE: 2000-10-05
; PRIOR APPLICATION NUMBER: 60/158,561
; PRIOR FILING DATE: 1999-10-08
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15
; LENGTH: 125
; TYPE: PRF
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(125)
; OTHER INFORMATION: LTB variant from NCBI gene bank GI: 3062900.
US-10-110-364-15

Query Match 100.0%; Score 106; DB 6; Length 125;
Best Local Similarity 100.0%; Pred. No. 6.6e-11;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GATFOVEVPGSOHIDSOKKAI 21
DB 67 GATFOVEVPGSOHIDSOKKAI 87

RESULT 11
US-10-141-627-2
; Sequence 2, Application US/10141627
; GENERAL INFORMATION:
; APPLICANT: Dale, James B.
; TITLE OF INVENTION: ANTIGEN OF HYBRID M PROTEIN AND CARRIER
; FILE REFERENCE: 481112.404C3
; CURRENT APPLICATION NUMBER: US/10/141,627
; CURRENT FILING DATE: 2002-05-07
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 138
; TYPE: PRF
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: LT-B-M24 hybrid molecule
US-10-141-627-2

Query Match 100.0%; Score 106; DB 6; Length 138;
Best Local Similarity 100.0%; Pred. No. 7.4e-11;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GATFOVEVPGSOHIDSOKKAI 21
DB 66 GATFOVEVPGSOHIDSOKKAI 86

RESULT 12
US-10-110-364-22
; Sequence 22, Application US/10110364
; GENERAL INFORMATION:

; APPLICANT: Handley, Harold H.
; APPLICANT: Haaparanta, Tapio
; APPLICANT: Ewalt, Karla L.
; TITLE OF INVENTION: ABS TOXIN B SUBUNIT MUTANTS WITH ALTERED
; FILE REFERENCE: ACTBIO.004A
; CURRENT APPLICATION NUMBER: US/10/110,364
; CURRENT FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: PCT/US00/27607
; PRIOR FILING DATE: 2000-10-05
; PRIOR APPLICATION NUMBER: 60/158,561
; PRIOR FILING DATE: 1999-10-08
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 22
; LENGTH: 103
; TYPE: PRF
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(103)
; OTHER INFORMATION: LTB variant from NCBI gene bank GI: 494265.
US-10-110-364-22

Query Match 95.3%; Score 101; DB 6; Length 103;
Best Local Similarity 95.2%; Pred. No. 3.9e-10;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GATFOVEVPGSOHIDSOKKAI 21
DB 45 GATFOVEVPGSOHIDSOKKAI 65

RESULT 13
US-10-110-364-20
; Sequence 20, Application US/10110364
; GENERAL INFORMATION:
; APPLICANT: Handley, Harold H.
; APPLICANT: Haaparanta, Tapio
; APPLICANT: Ewalt, Karla L.
; TITLE OF INVENTION: ABS TOXIN B SUBUNIT MUTANTS WITH ALTERED
; FILE REFERENCE: ACTBIO.004A
; CURRENT APPLICATION NUMBER: US/10/110,364
; CURRENT FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: PCT/US00/27607
; PRIOR FILING DATE: 2000-10-05
; PRIOR APPLICATION NUMBER: 60/158,561
; PRIOR FILING DATE: 1999-10-08
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 20
; LENGTH: 123
; TYPE: PRF
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(123)
; OTHER INFORMATION: LTB variant from NCBI gene bank GI: 223254.
US-10-110-364-20

Query Match 95.3%; Score 101; DB 6; Length 123;
Best Local Similarity 95.2%; Pred. No. 4.8e-10;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GATFOVEVPGSOHIDSOKKAI 21
DB 65 GATFOVEVPGSOHIDSOKKAI 85

RESULT 14

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 3, 2002, 08:56:32 ; Search time 46.57 Seconds
(without alignments)
43.330 Million cell updates/sec

Title: US-09-786-648-4

Sequence: 106
1 GATFOVEVPGSOHDSQKKAI 21

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	101	95.3	124	1 QLECB	heat-labile entero
2	100	94.3	124	1 XVVCB	cholera enterotoxi
3	48	45.3	392	2 T04150	RAD23 protein homo
4	46	43.4	374	2 T19866	hypothetical prote
5	46	43.4	574	2 C86400	hypothetical prote
6	44.5	42.0	374	2 D81715	conserved hypotnet
7	44.5	42.0	650	2 AG0733	probable bacterioph
8	44	41.5	91	2 C96580	hypothetical prote
9	44	41.5	255	2 A86457	probable peptide c
10	44	41.5	1742	2 S76110	hypothetical prote
11	43	40.6	365	2 H69231	sensory transducti
12	43	40.6	399	2 T46898	queuine tRNA-ribos
13	43	40.6	439	2 T49189	kinesin heavy chai
14	43	40.6	935	2 AE2469	two-component hybr
15	42.5	40.1	376	1 S17246	chorismate synthas
16	42	39.6	216	2 F83962	hypothetical prote
17	42	39.6	299	2 JG0178	chitinase (EC 3.2.
18	42	39.6	427	2 JG5694	stress-activated p
19	42	39.6	787	2 G81692	inner membrane pro
20	42	39.6	864	2 T49574	probable carnitine
21	41.5	39.2	500	2 JG4022	4-aminobutyrate tr
22	41.5	39.2	1090	2 S11823	alpha-dextrin endo
23	41	38.7	91	2 F81109	conserved hypotnet
24	41	38.7	128	1 UQDOR	ubiquitin / riboso
25	41	38.7	154	1 UQDOR7	ubiquitin / riboso
26	41	38.7	228	2 D34080	ubiquitin 18 - sli
27	41	38.7	229	2 B27806	ubiquitin (clone 1
28	41	38.7	313	2 E88216	protein B0495.8 [i
29	41	38.7	380	2 C34080	polyubiquitin 5 (c

30	41	38.7	380	2 B34080	polyubiquitin 5 (c
31	41	38.7	381	2 A27806	polyubiquitin 5 (c
32	41	38.7	386	1 S66056	yeast protein - Bac
33	41	38.7	532	2 A34080	polyubiquitin 7 (c
34	41	38.7	550	2 T01770	hypothetical prote
35	41	38.7	2911	2 T20566	hypothetical prote
36	40.5	38.2	862	2 S47598	multi protein homol
37	40	37.7	79	2 C95152	1566 family elemen
38	40	37.7	79	2 H95167	1566 family elemen
39	40	37.7	105	2 G69903	hypothetical prote
40	40	37.7	227	2 T32894	hypothetical prote
41	40	37.7	242	2 T34767	hypothetical prote
42	40	37.7	260	2 C96827	protein F20B17.2 [
43	40	37.7	330	2 C89848	peptide chain rele
44	40	37.7	355	2 T24938	hypothetical prote
45	40	37.7	367	2 T01751	gibberellin 20-oxi

ALIGNMENTS

RESULT 1
QLECB
heat-labile enterotoxin chain B precursor - Escherichia coli
C:Species: Escherichia coli
C:Date: 29-Jun-1981 #sequence_revision 29-Jun-1981 #text_Change 18-Jun-1999
C:Accession: A01820; B26946; I41194; I41287; I67644; A61475
R:Dallas, W.S.; Falkow, S.
Nature 288, 499-501, 1980
A:Title: Amino acid sequence homology between cholera toxin and Escherichia coli heat
A:Reference number: A01820; MUID:81074965
A:Accession: A01820
A:Molecule type: mRNA
A:Residues: 1-124 <DAL>
R:Yamamoto, T.; Gojobori, T.; Yokota, T.
J. Bacteriol. 169, 1352-1357, 1987
A:Title: Evolutionary origin of pathogenic determinants in enterotoxigenic Escherich
A:Reference number: A26946; MUID:87137303
A:Accession: B26946
A:Molecule type: DNA
A:Residues: 1-27, 'E', '29-63, 'K', '65-124 <YAM>
A:Cross-references: EMBL:M1363; NID:g148335; PIDN:AAA24792.1; PID:g148336
R:Leong, J.; Vinal, A.C.; Dallas, W.S.
Infect. Immun. 48, 73-77, 1985
A:Title: Nucleotide sequence comparison between heat-labile toxin B-subunit cistrons
A:Reference number: I41194; MUID:85156481
A:Accession: I41194
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-5, 'E', '7-17, 'C', '19-24, 'S', '26-27, 'E', '29-33, 'H', '35-63, 'K', '65-66, 'A', '68-122
A:Cross-references: GB:M17874; NID:g145830; PIDN:AAA98064.1; PID:g145831
A:Experimental source: plasmid ENT-R PCG86
R:Ibrahim, I.; Gentz, R.
J. Biol. Chem. 262, 10189-10194, 1987
A:Title: A functional interaction between the signal peptide and the translation appa
liculum.
A:Reference number: I41287; MUID:87280041
A:Accession: I41287
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-22 <RE2>
A:Cross-references: GB:M17101; NID:g146375; PIDN:AAA23973.1; PID:g146376
R:Imoue, T.; Tsuji, T.; Koto, M.; Imanura, S.; Miyama, A.
FEBS Microbiol. Lett. 108, 157-161, 1993
A:Title: Amino acid sequence of heat-labile enterotoxin from chicken enterotoxigenic
A:Reference number: I53542; MUID:99252225
A:Accession: I67644
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-17, 'C', '19-24, 'S', '26-27, 'E', '29-63, 'K', '65-66, 'A', '68-122, 'E', '124 <R
A:Cross-references: GB:560731; NID:g408994; PIDN:AA60441.1; PID:g408996
R:Tsuji, T.; Iida, T.; Honda, T.; Mawatani, T.; Nagahama, M.; Sakurai, J.; Wada, K.;
Microb. Pathog. 2, 381-390, 1987

A:Title: A unique amino acid sequence of the B subunit of a heat-labile enterotoxin iso
A:Reference number: A61475; MUID:89180953
A:Accession: A61475
A:Molecule type: protein
A:Residues: 22-24,'S',26-27,'E',29-63,'K',65-66,'A',68-95,'A',97-122,'E',124 <ISU>
A:Experimental source: strain 240-3
C:Complex: the heat-labile enterotoxin molecule contains one A chain and five or six B c
C:Function:
A:Description: the biological activity of the toxin is produced by the A chain, which ac
C:Superfamily: cholera enterotoxin beta chain
C:Keywords: enterotoxin
F:1-21/Domain: signal sequence #status predicted <SIG>
F:22-124/Product: heat-labile enterotoxin chain B #status predicted <MAT>
F:30-107/Disulfide bonds: #status predicted

Query Match 95.3%; Score 101; DB 1; Length 124;
Best Local Similarity 95.2%; Pred. No. 2.9e-09;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GATFOVEVPGSQHIDSQKRAI 21
Db 66 GEFQVEVPGSQHIDSQKRAI 86

RESULT 2
XVVCB
Cholera enterotoxin, B chain precursor VC1456 [validated] - Vibrio cholerae (strain N169
N:Alternate names: enterotoxin beta chain
C:Species: Vibrio cholerae
C:Date: 24-Apr-1984 #sequence,revision 01-Sep-2000 #text-change 02-Feb-2001
C:Accession: S14624; S39238; S39241; H62196; JC1078; S17666; PC1010; A05130; A01819; A38
R:Dams, E.; de Wolf, M.; Dierick, W.
submitted to the EMBL Data Library, March 1991
A:Description: Correction of the cholera toxin nucleotide sequence of the Vibrio cholera
A:Reference number: S14623
A:Accession: S14624
A:Molecule type: DNA
A:Residues: 1-124 <DAM>
A:Cross-references: EMBL:X58786; NID:g48420; PIDN:CAA41593.1; PID:g48422
A:Experimental source: strain 2125
R:Lebens, M.; Holmgren, J.
submitted to the EMBL Data Library, November 1993
A:Description: Structure and arrangement of the Cholera toxin genes in vibrio Cholera 01
A:Reference number: S39238
A:Accession: S39238
A:Molecule type: DNA
A:Residues: 1-124 <LEB>
A:Cross-references: EMBL:X76390; NID:g433856; PIDN:CAA53973.1; PID:g433857
A:Accession: S39241
A:Molecule type: DNA
A:Residues: 1-124 <LEN>
A:Cross-references: EMBL:X76391; NID:g433859; PIDN:CAA53976.1; PID:g433861
R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwin, M.L.; Dodson, R.J.;
Chardson, D.; Esmolaeva, M.D.; Yamahevan, J.; Bess, S.; Qin, H.; Dragol, I.; Sellers, F
L. R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A:Reference number: A82035; MUID:20406833
A:Accession: H82196
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-124 <HEI>
A:Cross-references: GB:AE004224; GB:AE003852; NID:g9655952; PIDN:AAF94613.1; GSPDB:GN001
A:Experimental source: serogroup O1, strain N16961; biotype El Tor
R:Shi, C.H.; Cao, C.; Zhang, J.S.; Ma, Q.J.
Chinese Biochem. J. 9, 395-399, 1993
A:Title: Nucleotide sequence analysis of the gene encoding the classical biotype V.chol
A:Reference number: JC1078
A:Accession: JC1078
A:Molecule type: DNA
A:Residues: 1-20,'O',22-31,'O',33-38,'H',40-49,'G',51-67,'T',69-124 <SHI>
A:Experimental source: classical biotype strain 569B

R:Dams, E.; de Wolf, M.; Dierick, W.
Biochem. Biophys. Acta 1090, 139-141, 1991
A:Title: Nucleotide sequence analysis of the CT operon of the Vibrio cholerae classic
A:Reference number: S17665; MUID:91355224
A:Accession: S17666
A:Molecule type: DNA
A:Residues: 1-38,'H',40-67,'T',69-124 <DA2>
A:Cross-references: EMBL:X58785; NID:g48888; PIDN:CAA41591.1; PID:g48890
R:Ma, Q.J.; Liu, C.X.; Xiong, L.S.; Yu, X.Q.
Sci. Sin. B Chem. Biol. Agric. Med. Earth Sci. 34, 274-280, 1991
A:Title: B subunit of cholera toxin produced in Escherichia coli.
A:Reference number: PC1010
A:Accession: PC1010
A:Molecule type: protein
A:Residues: 22-38,'H',40-41 <MAQ>
R:Mekalanos, J.J.; Swartz, D.J.; Pearson, G.D.N.; Harford, N.; Groyne, F.; de Wilde,
Nature 306, 551-557, 1983
A:Reference number: A93320; MUID:84068199
A:Accession: A05130
A:Molecule type: DNA
A:Residues: 1-32,'S',34-74,'S',76-124 <MEK>
A:Cross-references: GB:X00171; NID:g48347; PIDN:CAA24996.1; PID:g758351
R:Kurosky, A.; Markel, D.E.; Peterson, J.W.
J. Biol. Chem. 252, 7257-7264, 1977
A:Title: Covalent structure of the beta chain of cholera enterotoxin.
A:Reference number: A01819; MUID:7800537
A:Accession: A01819
A:Molecule type: protein
A:Residues: 22-38,'H',40-42,'N',44-67,'T',69-90,'N',92-124 <KUR>
R:Ali, C.Y.
J. Biol. Chem. 252, 7249-7256, 1977
A:Title: Determination of the primary structure of cholera toxin B subunit.
A:Reference number: A38033; MUID:7800536
A:Accession: A38033
A:Molecule type: protein
A:Residues: 22-38,'H',40-42,'N',44-67,'T',69,'E',71,'Q',74-75,'VE',78-86,'Q',101-103
A:Note: the difference at residue 70 may be due to deamidation during preparation
R:Nakashima, Y.; Napierkowski, P.; Schafer, D.E.; Konigsberg, W.H.
FEBS Lett. 68, 275-278, 1976
A:Title: Primary structure of the B subunit of cholera enterotoxin.
A:Reference number: A38034; MUID:77026365
A:Accession: A38034
A:Molecule type: protein
A:Residues: 22-38,'H',40-67,'T',69,'E',71,'Q',74-75,'VE',78-86,'Q',101-103
R:Takao, T.; Watanabe, H.; Shimomishi, Y.
Eur. J. Biochem. 146, 503-508, 1985
A:Title: Facile identification of protein sequences by mass spectrometry.
A:Reference number: A21910; MUID:85126976
A:Accession: A21910
A:Molecule type: protein
A:Residues: 22-38,'H',40-42,'N',44-67,'T',69-90,'N',92-124 <TAK>
A:Experimental source: biotype Inaba 569B
A:Note: Asn-65 was partially deaminated to Asp.
C:Comment: The authors translated the codon TCA for residue 33 as Tyr.
C:Genetics:
A:Gene: VC1456
A:Map position: 1
C:Complex: the cholera enterotoxin molecule contains three kinds of chains; an alpha
cate noncovalently with the subunit B, an aggregate of five beta chains
C:Function:
A:Description: involved in binding of the toxin to cell membranes
C:Superfamily: cholera enterotoxin beta chain
C:Keywords: enterotoxin; toxin
F:1-21/Domain: signal sequence #status predicted <SIG>
F:22-124/Product: cholera enterotoxin chain B #status predicted <MAT>
F:30-107/Disulfide bonds: #status experimental

Query Match 94.3%; Score 100; DB 1; Length 124;
Best Local Similarity 95.2%; Pred. No. 4.3e-09;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GATFOVEVPGSQHIDSQKRAI 21

RESULT 5
 C86400
 hypothetical protein T17H3.2 - *Arabidopsis thaliana*
 C:Species: *Arabidopsis thaliana* (mouse-ear cress)
 C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001
 C:Accession: C86400
 R:Theologis, A.; Eckert, J.R.; Palm, C.J.; Federpiel, N.A.; Kaul, S.; White, O.; Alonso,
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Cressy, T.H.; Dewar, K.

```

RESULT      7
AG0733
probable bacteriophage protein STY2026 [imported] - Salmonella enterica subsp. enterica
C:Species: Salmonella enterica subsp. enterica serovar Typhi
A:Note: This species has also been called Salmonella typhi
C:Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 09-Nov-2001
C:Accession: AG0733
R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Church
th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar
, S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens,
A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica ser
A:Reference number: AB0502; PMID:11677608
A:Accession: AG0733
A:Status: preliminary

```

A:Molecule type: DNA
A:Residues: 1-650 <PAR>
A:Cross-references: GB:AL513382; PIDN:CAD05569.1; PID:g16503066; GSPDB:GN00176
C:Genetics:
A:Gene: STY2026

Query Match 42.0%; Score 44.5; DB 2; Length 650;
Best Local Similarity 37.5%; Pred. No. 35;
Matches 9; Conservative 5; Mismatches 7; Indels 3; Gaps 1;
QY 1 GATFOVEVPGSGHIDSOKKAI 21
DB 464 GANIQDIPGAEVPOHQAQSAKRTL 487

RESULT 8
C96580
hypothetical protein F1511.6 (imported) - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C:Accession: C96580
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maitl, R.; Marziani,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719
A:Accession: C96580
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-91 <STO>
A:Cross-references: GB:AE005173; NID:g4587538; PIDN:AAD35769.1; GSPDB:GN00141
C:Genetics:
A:Gene: F1511.6
A:Map position: 1

Query Match 41.5%; Score 44; DB 2; Length 91;
Best Local Similarity 47.6%; Pred. No. 5.3;
Matches 10; Conservative 2; Mismatches 9; Indels 0; Gaps 0;

QY 1 GATFOVEVPGSGHIDSOKKAI 21
DB 58 GKTFFLEVKSGSETIQVKNMI 78

RESULT 9
A86457
probable peptide chain release factor F10C21.2 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 02-Nov-2001
C:Accession: A86457
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maitl, R.; Marziani,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719
A:Accession: A86457
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-255 <STO>

A:Cross-references: GB:AE005172; NID:g10645493; PIDN:AG21605.1; GSPDB:GN00141
C:Genetics:
A:Map position: 1

Query Match 41.5%; Score 44; DB 2; Length 255;
Best Local Similarity 42.1%; Pred. No. 16;
Matches 8; Conservative 4; Mismatches 7; Indels 0; Gaps 0;
QY 3 TFOVEVPGSGHIDSOKKAI 21
DB 102 TFRVSGFGGGRKRRKDSAV 120

RESULT 10
S76110
hypothetical protein - Synechocystis sp. (strain PCC 6803)
C:Species: Synechocystis sp.
A:Variety: PCC 6803
C:Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 08-Oct-1999
C:Accession: S76110
R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima,
O. K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yas
DNA Res. 3, 109-136, 1996
A:Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocys
s.
A:Reference number: S74322; MUID:97061201
A:Accession: S76110
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-1742 <KAN>
A:Cross-references: EMBL:D63999; GB:AB001339; NID:g1001396; PIDN:BA110088.1; PID:d101
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996
C:Genetics:
A:start codon: GTG

Query Match 41.5%; Score 44; DB 2; Length 1742;
Best Local Similarity 38.1%; Pred. No. 1.2e+02;
Matches 8; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

QY 1 GATFOVEVPGSGHIDSOKKAI 21
DB 121 GGTFSIDVPGSALVADPDRIT 141

RESULT 11
H69231
sensory transduction histidine kinase - Methanobacterium thermoautotrophicum (strain
C:Species: Methanobacterium thermoautotrophicum
C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 22-Oct-1999
C:Accession: H69231
R:Smith, D.R.; Doucette-Stamm, L.A.; Delonghery, C.; Lee, H.; Dubois, J.; Aldredge, T.
; Olu, D.; Spadafora, R.; Vitale, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jivani,
K.J.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.
J. Bacteriol. 179, 7135-7155, 1997
A:Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: fu
A:Reference number: A69000; MUID:98037514
A:Accession: H69231
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-365 <MTH>
A:Cross-references: GB:AE000872; GB:AE000666; NID:g2622082; PIDN:AAB85482.1; PID:g262
A:Experimental source: strain Delta H
C:Genetics:
A:Gene: MTH985

Query Match 40.6%; Score 43; DB 2; Length 365;
Best Local Similarity 36.8%; Pred. No. 34;
Matches 7; Conservative 6; Mismatches 6; Indels 0; Gaps 0;

QY 2 ATFOVEVPGSGHIDSOKKA 20

Mon Jul 8 07:51:55 2002

us-09-786-648-4.rpr

Page 6

Search completed: July 3, 2002, 08:56:33
Job time: 233 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 3, 2002, 09:04:42 ; Search time 21.51 seconds

(without alignments)
37.802 Million cell updates/sec

Title: US-09-786-648-4

Sequence: 1 GATGVEVPGSGHIDSQKAI 21

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	106	100.0	124	1	ELBH_ECOLI
2	101	95.3	124	1	ELBP_ECOLI
3	100	94.3	124	1	CHTB_VIBCH
4	44	41.5	467	1	IRF6_HUMAN
5	44	41.5	467	1	IRF6_MOUSE
6	43	40.6	376	1	TGT_ZYMO
7	42.5	40.1	376	1	AROC_YEAST
8	42	39.6	855	1	STI4_HUMAN
9	41.5	39.2	500	1	GABT_HUMAN
10	41.5	39.2	1090	1	PULA_KLEPN
11	41	38.7	76	1	UBIO_DICDI
12	41	38.7	313	1	YP68_CAEEL
13	41	38.7	386	1	YAAN_BACSU
14	40.5	38.2	862	1	PSM2_HUMAN
15	40	37.7	288	1	HS74_CAEEL
16	40	37.7	309	1	CLT2_RAT
17	40	37.7	355	1	YR1_CAEEL
18	40	37.7	655	1	YDH2_SCHPO
19	40	37.7	683	1	RPC_BPPHC
20	40	37.7	992	1	DLPI_RAT
21	40	37.7	2733	1	RRPH_CWMA5
22	39.5	37.3	560	1	INRI_SHEEP
23	39	36.8	112	1	Y13K_SSV1
24	39	36.8	263	1	YC43_ODOS1
25	39	36.8	456	1	SRS4_THEAC
26	39	36.8	459	1	IL7R_MOUSE
27	39	36.8	491	1	CD5_RAT
28	39	36.8	661	1	HS7C_CAEEL
29	39	36.8	664	1	CC4_CAVANL
30	39	36.8	704	1	PCCA_RAT
31	39	36.8	775	1	TH11_SCHPO
32	39	36.8	842	1	DPOL_HPEVM
33	39	36.8	1634	1	PK3B_HUMAN

34	39	36.8	2731	1	RRPB_CWMAH	P29982 murine coro
35	38.5	36.3	293	1	CHTA_CICAR	P36908 cicer ariet
36	38.5	36.3	344	1	LICH_PSESS	P25276 pseudomonas
37	38.5	36.3	529	1	TACY_LISMO	P13128 listeria mo
38	38.5	36.3	744	1	CATA_EMENTI	P55305 emeritella
39	38	35.8	76	1	UBIO_NEUCR	P13117 neuropept
40	38	35.8	148	1	IF5A_AERPE	O99453 aeropyrum p
41	38	35.8	169	1	Y459_MYCGE	O99436 mycoplasma
42	38	35.8	169	1	Y459_MYCPN	P75118 mycoplasma
43	38	35.8	205	1	HAP3_KIOLA	P40914 kluyveromyc
44	38	35.8	259	1	MOB2_YEAST	P43563 saccharomyc
45	38	35.8	359	1	THD1_SOLTU	P31212 solanum tub

ALIGNMENTS

RESULT	ID	STANDARD	PRT	AA
1	ELBH_ECOLI			
AC	P13811			
DT	01-JAN-1990 (Rel. 13, Created)			
DT	01-JAN-1990 (Rel. 13, last sequence update)			
DT	15-JUL-1999 (Rel. 38, last annotation update)			
DE	Heat-labile enterotoxin B chain precursor (LT-B, human) (LTB-B).			
GN	ELTB OR LTPB.			
OS	Escherichia coli.			
OC	Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;			
OC	Escherichia.			
OX	NCBI_TaxID=562;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-ISOLATE H74-114;			
RX	MEDLINE=85156481; PubMed=3884513;			
RA	Leong J., Vinal A.C., Dallas W.S.;			
RT	"Nucleotide sequence comparison between heat-labile toxin B-subunit			
RT	clstrons from Escherichia coli of human and porcine origin."			
RL	Infect. Immun. 48:73-77(1985).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-ISOLATE H10407;			
RX	MEDLINE=83114628; PubMed=6759877;			
RA	Yamamoto T., Tamura T.A., Yokota T., Takano T.;			
RT	"overlapping genes in the heat-labile enterotoxin operon originating			
RT	from Escherichia coli human strain."			
RL	Mol. Gen. Genet. 188:356-359(1982).			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-ISOLATE H10407;			
RX	MEDLINE=93252225; PubMed=8486242;			
RA	Inoue T., Tsuji T., Koto M., Imamura S., Miyama A.;			
RT	"Amino acid sequence of heat-labile enterotoxin from chicken			
RT	enterotoxigenic Escherichia coli is identical to that of human strain			
RL	H 10407."			
RL	FEMS Microbiol. Lett. 108:157-161(1993).			
RN	[4]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-ETEC LT 87;			
RA	Germain Y., Desperrier J.M.;			
RT	Submitted (Oct-1996) to the EMBL/GenBank/DBJ databases.			
RN	[5]			
RP	DISCUSSION OF SEQUENCE.			
RC	MEDLINE=95349400; PubMed=7623669;			
RA	Domenghini M., Pizsa M., Jobling M.G., Holmes R.K., Rappuoli R.;			
RT	"Identification of errors among database sequence entries and			
RT	comparison of correct amino acid sequences for the heat-labile			
RT	enterotoxins of Escherichia coli and Vibrio cholerae."			
RL	Mol. Microbiol. 15:1165-1167(1995).			
RN	[6]			
RP	X-RAY CRYSTALLOGRAPHY (3.04 ANGSTROMS).			
RC	MEDLINE=99185101; PubMed=10085117;			
RA	Matkovic-Calogovic D., Loregian A., D'Acunzio M.R., Battistutta R.,			
RT	Tossi A., Palu G., Zanotti G.;			

```
RT "Crystal structure of the B subunit of Escherichia coli heat-labile
RT enterotoxin carrying peptides with anti-herpes simplex virus type 1
RT activity."
RL J. Biol. Chem. 274:8764-8769(1999).
CC -1- FUNCTION: THE BIOLOGICAL ACTIVITY OF THE TOXIN IS PRODUCED BY
CC THE A CHAIN, WHICH ACTIVATES INTRACELLULAR ADENYL CYCLASE.
CC -1- SUBUNIT: HETEROHEXAMER OF ONE A CHAIN AND OF FIVE B CHAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M17874; AAA98064.1; -.
DR EMBL: J01646; AAB02982.1; -.
DR EMBL: S60731; AAC6041.1; -.
DR EMBL: X83966; CAA58800.1; -.
DR PDB: 1LTR; 23-MAR-99.
DR InterPro: IPR001835; Enterotoxin_B.
DR Pfam: PF01376; Enterotoxin_B.
DR PRINTS: PR00772; ENTEROTOXIN_B.
DR ProDom: PD012805; Enterotoxin_B; 1.
DR Enterotoxin; Signal; 3D-structure.
FT SIGNAL 1 21
FT CHAIN 22 124 HEAT-LABILE ENTEROTOXIN B CHAIN.
FT DISULFID 30 107
FT VARIANT 6 6 F -> C (IN ISOLATE H10407).
FT VARIANT 20 20 H -> Y (IN ISOLATE H10407).
FT VARIANT 34 34 H -> R (IN ISOLATE H10407).
SQ SEQUENCE 124 AA; 14027 MW; E9F7F7C7B9D3BC47 CRC64;

Query Match 100.0%; Score 106; DB 1; Length 124;
Best Local Similarity 100.0%; Pred. No. 2,1e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GATFOVEVPGSQHIDSOKKAI 21
DB 66 GATFOVEVPGSQHIDSOKKAI 86

RESULT 2
ELBP_ECOLI STANDARD: PRT; 124 AA.
AC P32890; P33768; P01557;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Heat-labile enterotoxin B chain precursor (LT-B, porcine) (LTP-B).
GN ELTB OR LTPB.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-ISOLATE P307;
RX MEDLINE=81074965; PubMed=7003397;
RA Dallas W.S., Falkow S.;
RT "Amino acid sequence homology between cholera toxin and Escherichia
RT coli heat-labile toxin."
RL Nature 288:499-501(1980).
RN [2]
RP REVISIONS TO 28 AND 64.
RC STRAIN-ISOLATE P307;
RX MEDLINE=85156481; PubMed=3884513;
RA Leong J., Vinal A.C., Dallas W.S.;
RT "Nucleotide sequence comparison between heat-labile toxin B-subunit
RT cistrons from Escherichia coli of human and porcine origin."
RL Infect. Immun. 48:73-77(1985).
```

```
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-ISOLATE PC86;
RX MEDLINE=87137303; PubMed=3546273;
RA Yamamoto T., Gojibori T., Yokota T.;
RT "Evolutionary origin of pathogenic determinants in enterotoxigenic
RT Escherichia coli and Vibrio cholerae O1."
RL J. Bacteriol. 169:1352-1357(1987).
RN [4]
RP SEQUENCE OF 1-22 FROM N.A.
RX MEDLINE=87280041; PubMed=3301830;
RA Ibrahim I., Gentz R.;
RT "A functional interaction between the signal peptide and the
RT translation apparatus is detected by the use of a single point
RT mutation which blocks translocation across mammalian endoplasmic
RT reticulum."
RL J. Biol. Chem. 262:10189-10194(1987).
RN [5]
RP X-RAY CRYSTALLOGRAPHY (1.95 ANGSTROMS).
RX MEDLINE=93240541; PubMed=8478941;
RA Sixma T.K., van Zanten B.A.M., Dauter Z., Hol W.G.J.;
RT "Refined structure of Escherichia coli heat-labile enterotoxin, a
RT close relative of cholera toxin."
RL J. Mol. Biol. 230:890-918(1993).
RN [6]
RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS).
RX MEDLINE=91238966; PubMed=2034287;
RA Sixma T.K., Pronk S.E., Kalk K.H., Martna E.S., van Zanten B.A.M.,
RA Witolt B., Hol W.G.J.;
RT "Crystal structure of a cholera toxin-related heat-labile enterotoxin
RT from E. coli."
RL Nature 351:371-377(1991).
RN [7]
RP DISCUSSION OF SEQUENCE.
RX MEDLINE=95349400; PubMed=7623669;
RA Domenighini M., Pizsa M., Jobling M.G., Holmes R.K., Rappuoli R.;
RT "Identification of errors among database sequence entries and
RT comparison of correct amino acid sequences for the heat-labile
RT enterotoxins of Escherichia coli and Vibrio cholerae."
RL Mol. Microbiol. 15:1165-1167(1995).
CC -1- FUNCTION: THE BIOLOGICAL ACTIVITY OF THE TOXIN IS PRODUCED BY
CC THE A CHAIN, WHICH ACTIVATES INTRACELLULAR ADENYL CYCLASE.
CC -1- SUBUNIT: HETEROHEXAMER OF ONE A CHAIN AND OF FIVE B CHAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M17873; AAA98065.1; -.
DR EMBL: M15363; AAA24792.1; -.
DR EMBL: M17101; AAA23973.1; -.
DR PIR: A01820; QLECB.
DR PIR: B26946; QLECEB.
DR PDB: 1LTA; 31-JAN-94.
DR PDB: 1LTB; 31-JAN-94.
DR PDB: 1LTG; 15-SEP-95.
DR PDB: 1LTI; 17-AUG-96.
DR PDB: 1LTT; 31-JAN-94.
DR PDB: 1LTT; 31-JAN-94.
DR PDB: 1LTT; 07-JUL-97.
DR PDB: 1LTT; 16-JUN-97.
DR PDB: 1LTT; 03-DEC-97.
DR PDB: 1LTT; 03-DEC-97.
DR PDB: 1LTT; 20-APR-95.
DR PDB: 1LTT; 20-APR-95.
DR InterPro: IPR001835; Enterotoxin_B.
DR Pfam: PF01376; Enterotoxin_B; 1.
DR PRINTS: PR00772; ENTEROTOXIN_B.
DR ProDom: PD012805; Enterotoxin_B; 1.
KW Enterotoxin; Signal; 3D-structure.
```

FT SIGNAL 1 21
 FT CHAIN 22 124
 FT DISULFID 30 107
 FT HELIX 26 30
 FT TURN 31 32
 FT STRAND 36 43
 FT TURN 47 51
 FT STRAND 54 55
 FT TURN 58 62
 FT STRAND 64 65
 FT TURN 68 71
 FT STRAND 76 77
 FT HELIX 80 98
 FT TURN 99 100
 FT STRAND 103 109
 FT STRAND 115 123
 SQ SEQUENCE 124 AA; 14133 MW; 6DB/DE58395EAY0D CRC64;

Query Match 95.3%; Score 101; DB 1; Length 124;
 Best Local Similarity 95.2%; Pred. NO. 1.4e-09;
 Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GARFQVEVPSQHIHSOKKAI 21
 ID CHTB_VIBCH STANDARD; PRT; 124 AA.
 AC P01556: Q9J002;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 13-AUG-1987 (Rel. 05, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Cholera enterotoxin, beta chain precursor.
 GN CTXB OR TOXB OR VCI456.
 OS Vibrio cholerae.
 OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
 OX NCBI_TaxID=666;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=84061784; PubMed=6315707;
 RA Lockman H., Kaper J.B.;
 RT "Nucleotide sequence analysis of the A2 and B subunits of Vibrio
 cholerae enterotoxin.";
 RL J. Biol. Chem. 258:13722-13726(1983).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=EL TOR 2125;
 RX MEDLINE=84068199; PubMed=6646234;
 RA Mekalanos J.J., Swartz D.J., Pearson G.D.N., Harford N., Groyne F.,
 de Wilde M.;
 RT "Cholera toxin genes: nucleotide sequence, deletion analysis and
 RT vaccine development.";
 RL Nature 306:551-557(1983).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=EL TOR 2125;
 RA Dams E., de Wolf M., Dierick W.;
 RL Submitted (MAY-1991) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=4260B / SEROTYPE O139;
 RX MEDLINE=94237453; PubMed=8181723;
 RA Lebens M., Holmgren J.;
 RT "Structure and arrangement of the cholera toxin genes in Vibrio
 cholerae O139.";
 RL FEMS Microbiol. Lett. 117:197-202(1994).
 RN [5]
 RP SEQUENCE FROM N.A.
 RC STRAIN=1854 / O139-BENGAL;
 RA Yamamoto K., Do V.G.R.F., Xu M., Iida T., Miyatani T., Albert M.J.,

RA Honda T.;
 RL Submitted (MAY-1994) to the EMBL/GenBank/DBJ databases.
 RN [6]
 RP SEQUENCE FROM N.A.
 RC STRAIN=EL TOR N16961 / SEROTYPE O1;
 RX MEDLINE=20406833; PubMed=10952301;
 RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwinn M.L.,
 Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,
 Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
 Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,
 McDonald L., Utterback T., Fleischmann R.D., Nierman W.C., White O.,
 Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
 Fraser C.M.;
 RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio
 cholerae.";
 RL Nature 406:477-483(2000).
 RN [7]
 RP SEQUENCE OF 22-124.
 RX MEDLINE=78005537; PubMed=903363;
 RA Kurosky A., Markel D.E., Peterson J.W.;
 RT "Covalent structure of the beta chain of cholera enterotoxin.";
 RL J. Biol. Chem. 252:7257-7264(1977).
 RN [8]
 RP SEQUENCE OF 22-124.
 RX MEDLINE=78005536; PubMed=903362;
 RA Lai C.-Y.;
 RT "Determination of the primary structure of cholera toxin B subunit.";
 RL J. Biol. Chem. 252:7249-7256(1977).
 RN [9]
 RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS).
 RX MEDLINE=94272319; PubMed=8003954;
 RA Merritt E.A., Sarfaty S., van den Akker F., L'Hoir C., Martial J.A.,
 Hol W.G.J.;
 RT "Crystal structure of cholera toxin B-pentamer bound to receptor GM1
 pentasaccharide.";
 RL Protein Sci. 3:166-175(1994).
 RN [10]
 RP X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS).
 RX MEDLINE=95387394; PubMed=7658472;
 RA Zhang R.G., Westbrook M.L., Westbrook E.M., Scott D.L., Otwinowski Z.,
 Maulik P.R., Reed R.A., Shipley G.G.;
 RT "The 2.4 A crystal structure of cholera toxin B subunit pentamer:
 RT choleragenoid.";
 RL J. Mol. Biol. 251:550-562(1995).
 RN [11]
 RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).
 RC STRAIN=OGAWA 41 / CLASSICAL BIOTYPE;
 RX MEDLINE=97376625; PubMed=9232653;
 RA Merritt E.A., Sarfaty S., Jobling M.G., Chang T., Holmes R.K.,
 Hirst T.R., Hol W.G.J.;
 RT "Structural studies of receptor binding by cholera toxin mutants.";
 RL Protein Sci. 6:1516-1528(1997).
 CC -1- FUNCTION: THE BETA CHAIN AGGREGATE (B SUBUNIT) IS INVOLVED IN
 BINDING TO CELL MEMBRANES.
 CC -1- SUBUNIT: CONTAINS 3 KINDS OF CHAINS, AN ALPHA AND A GAMMA CHAIN
 (FROM THE SAME PRECURSOR MOLECULE), LINKED BY AN INTERCHAIN
 DISULFIDE BOND, ASSOCIATE NONCOVALENTLY WITH AN AGGREGATE OF 4 TO
 6 BETA CHAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: X00171; CAA24996.1; -
 CC EMBL: X01170; AAA27573.1; -
 CC EMBL: D30053; BAA06291.1; -
 CC EMBL: X58786; CAA41593.1; -
 CC EMBL: X76390; CAA53973.1; -
 CC EMBL: X76391; CAA53976.1; -

```

DR EMBL: AE004224; AAF94613.1; -.
DR PIR: A01819; XVCB.
DR PIR: A05130; A05130.
DR PIR: S14624; S14624.
DR PDB: 2CHB; 03-DEC-97.
DR PDB: 3CHB; 12-AUG-98.
DR PDB: 1CHP; 08-MAR-96.
DR PDB: 1CHQ; 08-MAR-96.
DR PDB: 1FGB; 23-DEC-96.
DR PDB: 1XPB; 01-APR-97.
DR PDB: 1XTC; 01-AUG-96.
DR PDB: 1CT1; 15-OCT-97.
DR TIGR: VC1456; -.
DR InterPro: IPR001835; Enterotoxin_B.
DR Pfam: PF01376; Enterotoxin_B; 1.
DR PRINTS: PR00772; ENTEROTOXINB.
DR ProDom: PD012805; Enterotoxin_B; 1.
DR Membrane; Enterotoxin; Signal; 3D-structure; Complete proteome.
KW SIGNAL
FT CHAIN 1 21
FT DISULFID 22 124 CHOLERA ENTEROTOXIN, BETA CHAIN.
FT CONFLICT 30 107
FT CONFLICT 33 33 Y -> S (IN REF. 2).
FT CONFLICT 39 39 Y -> H (IN REF. 7 AND 8).
FT CONFLICT 43 43 D -> N (IN REF. 7 AND 8).
FT CONFLICT 68 68 I -> T (IN REF. 7 AND 8).
FT CONFLICT 70 70 Q -> E (IN REF. 8).
FT CONFLICT 75 75 G -> S (IN REF. 2).
FT CONFLICT 91 91 D -> N (IN REF. 7 AND 8).
FT HELIX 26 30
FT TURN 31 32
FT TURN 34 35
FT STRAND 36 44
FT STRAND 47 51
FT TURN 54 55
FT STRAND 58 62
FT TURN 64 65
FT STRAND 68 71
FT TURN 76 77
FT TURN 80 99
FT HELIX 100 100
FT TURN 102 109
FT STRAND 115 123
FT STRAND 124 AA; 13957 MM; 9AA393E3EA8E3EBF CRC64;
SQ SEQUENCE

```

```

Query Match 94.3%; Score 100; DB 1; Length 124;
Best Local Similarity 95.2%; Pred. No. 2e-09;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

OY 1 GATFOVEVPGSHIDSOKKAI 21
DB 66 GAIFOVEVPGSHIDSOKKAI 86
RESULT 4
ID IRF6_HUMAN STANDARD; PRT; 467 AA.
AC O14896;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Interferon regulatory factor 6 (IRF-6).
GN IRF6.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Grossman A., Mitrucker H.W., Antonio L., Ozato K., Mak T.W.;
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.

```

```

RA Grafton D.;
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: Nuclear (Potential).
CC -1- SIMILARITY: BELONGS TO THE IRF FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL: AF027292; AAB84111.1; -.
DR EMBL: AL022398; CAA18545.1; -.
DR HSSP: P23906; 2IRF.
DR InterPro: IPR001346; IRF.
DR Pfam: PF00605; IRF; 1.
DR PRINTS: PR00267; INTERNREGCT.
DR ProDom: PD002355; IRF; 1.
DR SMART; SM00348; IRF; 1.
DR PROSITE; PS00601; IRF; 1.
KW Transcription regulation; DNA-binding; Nuclear protein.
FT DNA_BIND 9 111 TRYPROPHAN PENTAD REPEAT.
FT SEQUENCE 467 AA; 53129 MM; 7E28F5E0F5BA4053 CRC64;
SQ SEQUENCE

```

```

Query Match 41.5%; Score 44; DB 1; Length 467;
Best Local Similarity 41.2%; Pred. No. 12;
Matches 7; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

```

```

OY 5 OVEVPGSHIDSOKKAI 21
DB 273 OVEVPGSHIDSOKKAI 289

```

```

RESULT 5
ID IRF6_MOUSE STANDARD; PRT; 467 AA.
AC P97431;
DT 01-MAR-2002 (Rel. 41, Created)
DT 01-MAR-2002 (Rel. 41, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Interferon regulatory factor 6 (IRF-6).
GN IRF6.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Grossman A., Mitrucker H.W., Antonio L., Mak T.W.;
RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: Nuclear (Potential).
CC -1- SIMILARITY: BELONGS TO THE IRF FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL: U73029; AAB36714.1; -.
DR HSSP: P23906; 2IRF.
DR MGD; MGI:1859211; Irf6.
DR InterPro: IPR001346; IRF.
DR Pfam; PF00605; IRF; 1.
DR PRINTS; PR00267; INTERNREGCT.
DR ProDom; PD002355; IRF; 1.
DR SMART; SM00348; IRF; 1.

```

DR PROSITE: PS00601; TRF: 1.
KW Transcription regulation; DNA-binding; Nuclear protein
FT DNA_BIND 9 111 TRYPTOPHAN PENTAD REPEAT.
SQ SEQUENCE 467 AA; 53106 MW; 68CCAA9680BEDC8 CRC64;

Query Match	41.5%	Score 44	DB 1	Length 467
Best Local Similarity	41.2%	Pred. No. 12		
Matches 7, Conservative	6	Mismatches 4	Indels 0	Gaps 0

```

OY      5 QVEVPGSQHIDSQKAI 21
         ||: ||: ||: ||: :
Db      273 QVKFPGEHITNEKQL 285

```

RESULT	6
ID	TGT_ZYMO
AC	P28720; STANDARD; PRT; 385 AA.
DT	01-DEC-1992 (Rel. 24, Created)
DT	01-NOV-1997 (Rel. 35, Last sequence update)
DT	01-MAR-2002 (Rel. 41, Last annotation update)
DE	Queuine tRNA-ribosyltransferase (EC 2.4.2.29) (tRNA-guanine transglycosylase) (Guanine insertion enzyme).
GN	TGT.
OS	Zymomonas mobilis.
OC	Bacteria; Proteobacteria; alpha subdivision; Sphingomonadaceae; Zymomonas.
OX	NCBI_Taxid=542;
RN	[1]
RP	SEQUENCE FROM N.A., SEQUENCE OF 1-6, AND CHARACTERIZATION.
RX	MEDLINE=95394847; Pubmed=7665516;
RA	Reuter K.K.H., Flicner R.;
RT	"Sequence analysis and overexpression of the Zymomonas mobilis tgt gene encoding tRNA-guanine transglycosylase: purification and biochemical characterization of the enzyme.";
RL	J. Bacteriol. 177:5284-5288(1995).
RN	[2]
RP	PARTIAL SEQUENCE FROM N.A.
RC	STRAIN=ATCC 31821 / ZM4 / CP4;
RX	MEDLINE=92406015 / Pubmed=15264662;
RA	Shark K.B., Conway T.;
RT	"Cloning and molecular characterization of the DNA ligase gene (lig) from Zymomonas mobilis.";
RL	FEMS Microbiol. Lett. 75:19-26(1992).
RN	[3]
RP	X-RAY CRYSTALLOGRAPHY (1.85 ANGSTROMS).
RX	MEDLINE=96256303; Pubmed=8654383;
RA	Romler C., Reuter K., Suck D., Flicner R.;
RT	"Crystal structure of tRNA-guanine transglycosylase: RNA modification by base exchange.";
RL	EMBO J. 15:2850-2857(1996).
RN	[4]
RP	X-RAY CRYSTALLOGRAPHY (1.95 ANGSTROMS).
RX	MEDLINE=21103718; Pubmed=11178905;
RA	Graedler U., Gaeber H.-D., Goodenough-Dashna D.M., Garcia G.A., Flicner R., Reuter K., Stubbs M.T., Klebe G.;
RT	"A new target for shiellois: rational design and crystallographic studies of inhibitors of tRNA-guanine transglycosylase.";
RL	J. Mol. Biol. 306:455-467(2001).
CC	-1- FUNCTION: Exchanges the guanine residue with 7-aminomethyl-7-deazaguanine in tRNAs with G(U) anticodons (tRNA-Asp, -Asn, -His and -Tyr). After this exchange, a cyclopentendiol moiety is attached to the 7-aminomethyl group of 7-deazaguanine, resulting in the hypermodified nucleoside queuine (Q) (7-((4,5-cis-dihydroxy-2-cyclopenten-1-yl)amino)methyl)-7-deazaguanosine).
CC	-1- CATALYTIC ACTIVITY: tRNA guanine + queuine -> tRNA queuine + guanine.
CC	-1- COFACTOR: BINDS AND REQUIRES ZINC FOR ACTIVITY. ALSO REQUIRES MAGNESIUM.
CC	-1- SUBUNIT: MONOMER.
CC	-1- SIMILARITY: BELONGS TO THE QUEUINE tRNA-RIBOSYLTRANSFERASE FAMILY.
CC	-1- CAUTION: REF. 2 SEQUENCE DIFFERS FROM THAT SHOWN EXTENSIVELY DUE

```

CC      TO FRAMESHIFTS.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as they are not used in any way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC      or send an email to license@sib-sib.ch).
CC      -----
DR      EMBL: L33777; AAA27704.1; ALT. INIT.
DR      EMBL: L33777; AAA27705.1; ALT. INIT.
DR      EMBL: 211910; -. NOT_ANNOTATED_CDS.
DR      PDB: 1PUD; 07-JUL-97.
DR      PDB: 1WKD; 07-JUL-97.
DR      PDB: 1WKE; 07-JUL-97.
DR      PDB: 1WKF; 07-JUL-97.
DR      PDB: 1ENU; 19-APR-00.
DR      InterPro: IPR002616; Que_rRNA-ribosyltransf.
DR      Pfam: PF01702; TGT; 1.
KM      Queuosine biosynthesis; Transferase; Glycosyltransferase;
KM      tRNA processing; Zinc; Magnesium; 3D-structure.
FT      INIT MET 0
FT      ACT SITE 102 102
FT      ACT SITE 280 280
FT      METAL 317 317 ZINC.
FT      METAL 319 319 ZINC.
FT      METAL 322 322 ZINC.
FT      METAL 348 348 ZINC.
SQ      SEQUENCE 385 AA; 42738 MW; F3D6FA270A0B2F3 CRC64;
OY      1 GATFQVEVPGSGHSDSOKKAI 21
OY      1 111:111:111:111:1
DB      120 GVTFFSHLDGSRHMLSPERSI 140

RESULT 7
AC      AROC_YEAST STANDARD: PRT; 376 AA.
AC      P28777;
DT      01-DEC-1992 (Rel. 24, Created)
DT      01-DEC-1992 (Rel. 24, Last sequence update)
DT      30-MAY-2000 (Rel. 39, Last annotation update)
DE      Chorismate synthase (EC 4.6.1.4) (5'-enolpyruvylshikimate-3-phosphate
DE      phosphorylase).
DE      GN ARO2 OR YGL148W.
OS      Saccharomyces cerevisiae (Baker's yeast).
OC      Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC      Saccharomycetales; Saccharomycetaceae; Saccharomyces.
NCBI_OX NCB1_maxID=4932;
RN      11
RP      SEQUENCE FROM N.A.
RC      STRAIN=S288C / FY1769;
RX      MEDLINE=92114793; PubMed=1837329;
RA      Jones D.G.L., Reusser U., Braus G.H.;
RA      "Molecular cloning, characterization and analysis of the regulation
RA      of the ARO2 gene, encoding chorismate synthase, of Saccharomyces
RA      cerevisiae.";
RL      Mol. Microbiol. 5:2143-2152(1991).
[2]
RP      SEQUENCE FROM N.A.
RC      STRAIN=S288C / FY1769;
RX      MEDLINE=92114793; PubMed=9046099;
RA      Voelt M., Defoor E., Verhasselt P., Riles L., Robben J., Volckaert G.;
RA      "The sequence of a nearly unclonable 22.8 kb segment on the left arm
RA      chromosome VII from Saccharomyces cerevisiae reveals ARO2, RPD9A,
RA      TIR1, MRF1 genes and six new open reading frames.";
RL      Yeast 13:177-182(1997).

```



```

KW Transmembrane; Repeat.
FT DOMAIN 1 55
FT TRANSMEM 56 76
FT
FT DOMAIN 77 85
FT DOMAIN 214 334
FT DOMAIN 340 447
FT DOMAIN 452 487
FT DOMAIN 487 524
FT DOMAIN 524 560
FT DOMAIN 566 603
FT DOMAIN 615 854
FT ACT_SITE 656 656
FT ACT_SITE 711 711
FT ACT_SITE 805 805
FT CARBOHYD 109 109
FT CARBOHYD 302 302
FT CARBOHYD 485 485
FT CARBOHYD 772 772
FT CONFLICT 327 329
FT CONFLICT 381 381
FT CONFLICT 674 674
SQ SEQUENCE 855 AA; 94769 MW; 26143132C01F96C9 CRC64;

Query Match 39.6%; Score 42; DB 1; Length 855;
Best Local Similarity 50.0%; Pred. No. 50;
Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

OY 3 TFOVEVPGSOH 14
Db 367 TWNTEVPMNOHV 378
1: :||| :||:

```

```

CC semialdehyde + L-glutamate.
CC -1- CORFACTOR: PYRIDOXAL PHOSPHATE.
CC -1- SUBUNIT: MONOMER (PROBABLE).
CC -1- SUBCELLULAR LOCATION: Mitochondrial matrix.
CC -1- TISSUE SPECIFICITY: LIVER > PANCREAS > BRAIN > KIDNEY > HEART > PLACENTA.
CC -1- DISEASE: DEFECTS IN ARAAT ARE A CAUSE OF GABA-AT DEFICIENCY WHOSE PHENOTYPE INCLUDES PSYCHOMOTOR RETARDATION, HYPOKONIA, HYPERREFLEXIA, LETHARGY, REFRACTORY SEIZURES, AND EEG ABNORMALITIES.
CC -1- SIMILARITY: BELONGS TO CLASS-III OF PYRIDOXAL-PHOSPHATE-DEPENDENT AMINOTRANSFERASES.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L32961; AAA74449.1; -.
DR HSSP; P80147; 1GTX.
DR MIM; 137150; -.
DR InterPro; IPR000954; Aminotran_3.
DR Pfam; PF00202; aminotran_3.1.
DR PROSITE; PS00600; AA_TRANSFER_CLASS_3; 1.
KW Transferrase; Aminotransferase; Pyridoxal phosphate; Neurotransmitter degradation; Mitochondrion; Transit peptide; Disease mutation.
KW TRANSIT 1 28
FT CHAIN 29 500
FT BINDING 357 357
FT VARIANT 220 220
FT
FT TRANSIT 1 28
FT CHAIN 29 500
FT BINDING 357 357
FT VARIANT 220 220
FT
FT CONFLICT 109 109
FT CONFLICT 113 113
FT CONFLICT 132 132
FT CONFLICT 191 191
FT CONFLICT 204 204
FT CONFLICT 216 216
FT CONFLICT 268 268
FT CONFLICT 320 320
FT CONFLICT 366 366
SQ SEQUENCE 500 AA; 56557 MW; 41199085693F80AD CRC64;

Query Match 39.2%; Score 41.5; DB 1; Length 500;
Best Local Similarity 55.6%; Pred. No. 34;
Matches 10; Conservative 4; Mismatches 3; Indels 1; Gaps 1;

OY 3 TFOVEVPGSOHDSOKA 20
Db 17 TYRLVPGSRHL-SQAAA 33
1::: ||||: || |

```

```

RESULT 10.
ID PULA_KLEPN STANDARD: PRT: 1090 AA.
AC P07206;
DT 01-APR-1988 (Rel. 07, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Pulinanase precursor (EC 3.2.1.41) (Alpha-dextrin endo-1,6-alpha-glucosidase) (Pulinan 6-glucanohydrolase).
GN PULA.
OS Klebsiella pneumoniae.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae; Klebsiella.
OX NCBI_TaxID=573;
RN [1]
RP SEQUENCE FROM N.A.

```

RC		STRAIN-UNF 5023;	
RX	MEDLINE=90205629; PubMed2181242;		
RA	Kornacker M.G., Pugsley A.P.;		
RT	"Molecular characterization of pulA and its product, pullulanase, a secreted enzyme of Klebsiella pneumoniae UNF5023.",		
RL	Mol. Microbiol. 4:73-85(1990).		
RN	(2)		
RP	SEQUENCE OF 1-62 FROM N.A.		
RX	MEDLINE=86033621; PubMed=3902792;		
RA	Chapon C., Raibaud O.;		
RT	"Structure of two divergent promoters located in front of the gene encoding pullulanase in Klebsiella pneumoniae and positively regulated by the malt product.";		
RL	J. Bacteriol. 164:639-645(1985).		
RN	[3]		
RP	SEQUENCE OF 944-1090 FROM N.A.		
RX	MEDLINE=89291709; PubMed=2661532;		
RA	D'Enfert C., Pugsley A.P.;		
RT	"Klebsiella pneumoniae puls gene encodes an outer membrane lipoprotein required for pullulanase secretion.";		
RL	J. Bacteriol. 171:3673-3679(1989).		
CC	-I- CATALYTIC ACTIVITY: Hydrolysis of (1->6)-alpha-D-glucosidic linkages in pullulan and in amylopectin and glycogen, and the alpha-and beta-limit dextrins of amylopectin and glycogen.		
CC	-I- SUBUNIT: HOMOTRIMER.		
CC	-I- SUBCELLULAR LOCATION: Attached to the membrane by a lipid anchor (probable).		
CC	-I- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO KNOWN AS THE ALPHA-AMYLASE FAMILY.		
CC	-- -- -- -- --		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isdb.ch/announce/ or send an email to license@isdb-slb.ch).		
CC	-- -- -- -- --		
DR	EMBL: X52181; CAA36431.1; -.		
DR	EMBL: M12503; AAA25087.2; -.		
DR	EMBL: M29097; AAA61976.1; -.		
DR	PIR: A25025; A25025.		
DR	PIR: A32880; A32880.		
DR	PIR: S11823; S11823.		
DR	InterPro: IPR0004631; Alpha_amylase. Interpro: IPR004193; isoamylase_N.		
DR	Pfam: PF00128; alpha-amylase_1.		
DR	PROSITE: PS00013; PROKARYOTEIN_LIPIDOLIPROTEIN. 1.		
KW	Hydrolase; Glycosidase; Membrane; Lipoprotein; Signal.		
FT	SIGNAL	1	19
FT	CHAIN	20	1090
FT	LIPID	20	20
FT	ACT_SITE	684	684
FT	ACT_SITE	713	713
FT	ACT_SITE	841	841
FT	CONFLICT	6	6
FT	CONFLICT	10	10
FT	CONFLICT	15	15
FT	CONFLICT	23	23
FT	CONFLICT	31	31
FT	CONFLICT	34	34
FT	CONFLICT	36	36
FT	CONFLICT	55	58
CQ	SEQUENCE	1090 AA:	118096 MW; 240AEVDFB3FFIBD6 CRC64;

```

Query Match          39.2%  Score 41.5:  DB 1;  Length 1090;
Best Local Similarity 38.5%:  Pred. No. 78;
Matches 10;  Conservative 4;  Mismatches 7;  Indels 5;  Gaps 1.
OY      1 GATFOVEVPGSOHID-----SOKKAI 21
      | | | | | | | : | | | : | | |

```

```
DB          3312      GVFPRVWAPTAQQVDVVYVVSADKVI   337
RESULT     11
UBIQ_DICDI STANDARD:             PRT:           76 AA.
AC    P08618:
DT    01-AUG-1988 (Rel. 08, Created)
DT    01-NOV-1988 (Rel. 08, Last sequence update)
DT    01-NOV-1991 (Rel. 20, Last annotation update)
DE    Ubiquitin.
OS    Dictyostelium discoideum (Slime mold).
OC    Eukaryota; Mycetozoa; Dictyostellida; Dictyostelium.
OX    NCBI_TaxID=44689;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=88152253; PubMed=2831095;
RA      Mueller-Taubenberger A., Westphal M., Jaeger E., Noegel A.,
RT      Gerisch G.;
RL      FEBS Lett. 229:273-278(1988).
RN      [2]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=89352609; PubMed=2548604;
RA      Omachi T., Giorda R., Shaw D.R., Ennis H.L.;
RT      "Molecular organization of developmentally regulated Dictyostelium
RT      discoidium ubiquitin cDNAs.";
RL      Biochemistry 28:5226-5231(1989).
RN      [3]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=87257921; PubMed=3037345;
RA      Giorda R., Ennis H.L.;
RT      "Structure of two developmentally regulated Dictyostelium discoidium
RT      ubiquitin genes.";
RL      Mol. Cell. Biol. 7:2097-2103(1987).
RN      [4]
RP      SEQUENCE OF 13-76 FROM N.A.
RC      STRAIN-AAx2;
RA      Westphal M., Mueller-Taubenberger A., Noegel A., Gerisch G.;
RT      "Transcript regulation and carboxyterminal extension of ubiquitin in
RT      Dictyostelium discoidium.";
RL      FEBS Lett. 209:52-96(1986).
CC      -I- FUNCTION: INVOLVED IN THE ATP-DEPENDENT SELECTIVE DEGRADATION OF
CC      CELLULAR PROTEINS, THE MAINTENANCE OF CHROMATIN STRUCTURE, THE
CC      REGULATION OF GENE EXPRESSION, THE STRESS RESPONSE, AND RIBOSOME
CC      BIOGENESIS.
CC      -I- SUBCELLULAR LOCATION: Nuclear and cytoplasmic.
CC      -I- MISCELLANEOUS: UBIQUITIN IS SYNTHESIZED AS A POLYUBIQUITIN
CC      PRECURSOR WITH EXACT HEAD TO TAIL REPEATS. SOME UBIQUITIN GENES
CC      CONTAIN A SINGLE COPY OF UBIQUITIN FUSED TO A RIBOSOMAL PROTEIN.
-----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
-----
DR      EMBL; X07210; CAA30183.1; ALT_TERM.
DR      EMBL; M19666; AAA33261.1; ALT_TERM.
DR      EMBL; M19491; AAA33269.1; ALT_TERM.
DR      EMBL; M19492; AAA33270.1; ALT_TERM.
DR      EMBL; M23748; AAA33262.1; ALT_TERM.
DR      EMBL; M23749; AAA33263.1; ALT_TERM.
DR      EMBL; M23750; AAA33264.1; ALT_TERM.
DR      EMBL; M23751; AAA33265.1; ALT_TERM.
DR      EMBL; M23752; AAA33266.1; ALT_TERM.
DR      EMBL; M23753; AAA33267.1; ALT_TERM.
DR      EMBL; M27544; AAA33268.1; ALT_TERM.
DR      EMBL; X04702; CAA28408.1; ALT_TERM.
```

DR PIR: S00357; UODOR.
DR PIR: E34080; UODOR7.
DR PIR: A27806; A27806.
DR PIR: B27806; B27806.
DR PIR: A34080; A34080.
DR PIR: B34080; B34080.
DR PIR: C34080; C34080.
DR PIR: D34080; D34080.
DR HSP: P02248; 1UBI.
DR DictyDb; DD05001; ubqA.
DR InterPro; IPR000626; ubiquitin.
DR Pfam; PF00240; ubiquitin.1.
DR PRINTS; PR00348; UBIQUITIN.
DR SMART; SM00213; UBO; 1.
DR PROSITE; PS00289; UBIQUITIN_1; 1.
DR PROSITE; PS50053; UBIQUITIN_2; 1.
KW Nuclear protein; Polyprotein.
FT SITE 48 48
FT BINDING 76 76
FT VARIANT 11 11
SQ SEQUENCE 76 AA; 8538 MW; 6427383968EAB84 CRC64;
NECESSARY FOR BRANCHED-CHAIN
MULTIUBIQUITIN ADDUCTS.
CONJUGATION TO ACCEPTOR PROTEINS.
K -> N (IN SOME CLONES REPEATS).

Query Match 38.7%; Score 41; DB 1; Length 76;
Best Local Similarity 42.9%; Pred. No. 5.3;
Matches 9; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

OY 1 GATFOVEVPGSHIDSOKKAI 21
10 GKTTLEVEGSDNIENKAKI 30

RESULT 12
ID YP68_CAEEL STANDARD; PRT; 313 AA.
AC 009217;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical 37.0 kDa protein B0495.8 in chromosome II.
GN B0495.8.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Pelodierinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RA [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RA Kirsten J.;
RL Submitted (FEB-1995) to the EMBL/GenBank/DBJ databases.
- i- SIMILARITY: SOME, TO YEAST YD1087C AND S.POMBE SPC16A11.13.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC
CC EMBL; U21317; AAA62527.1; -
DR WormPep; B0495.8; CE01766.
KW Hypothetical protein.
SQ SEQUENCE 313 AA; 36977 MW; 000D2327621BFED0 CRC64;

Query Match 38.7%; Score 41; DB 1; Length 313;
Best Local Similarity 46.2%; Pred. No. 25;
Matches 6; Conservative 6; Mismatches 1; Indels 0; Gaps 0;
OY 7 EVPGSHIDSOKK 19
:: |||||:::|

Db 11 QLMGSHVDNKK 23

RESULT 13
ID YAAN_BACSU STANDARD; PRT; 386 AA.
AC P37535;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical 43.8 kDa protein in XPAC-ABRB intergenic region.
GN YAAN.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Bacillus.
OX NCBI_TaxID=1423;
RA [1]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=96051385; PubMed=7584024;
RA Ogasawara N., Nakai S., Yoshikawa H.;
RT "Systematic sequencing of the 180 kilobase region of the Bacillus
RT subtilis chromosome containing the replication origin.";
RL DNA Res. 1:1-14(1994).
RN [2]
RP SEQUENCE OF 1-191 FROM N.A.
RA Bookstein C., Edwards C.W., Hulet F.M.;
RL Submitted (JUN-1992) to the EMBL/GenBank/DBJ databases.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC
CC EMBL; D26185; BAA05262.1; -
DR EMBL; M96156; AAA22892.1; -
DR EMBL; Z99104; CAB11802.1; -
KW Subtilisin; Bg10090; yaan.
DR Hypothetical protein; Complete proteome.
FT CONFLICT 191 191 E -> M (IN REF. 2).
SQ SEQUENCE 386 AA; 43830 MW; D921F3A0F6845EBE CRC64;

Query Match 38.7%; Score 41; DB 1; Length 386;
Best Local Similarity 31.2%; Pred. No. 31;
Matches 5; Conservative 8; Mismatches 3; Indels 0; Gaps 0;

OY 6 VEPGSHIDSOKKAI 21
22 IEIPGSEAVKAEKQOV 37

RESULT 14
ID PMS2_HUMAN STANDARD; PRT; 862 AA.
AC P54278;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE PMS1 protein homolog 2 (DNA mismatch repair protein PMS2).
GN PMS2 OR PMSL2.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RA [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Endometrial tumor;
RX MEDLINE=94352394; PubMed=8072530;
RA Nicolaides N.C., Papadopoulos N., Liu B., Wei Y.-F., Carter K.C.,

DR	EMBL	U13696	AAA63923.1	-	
DR	EMBL	U14658	AAA50390.1	-	
DR	HSSP	P23367	1BKN		
DR	SWISS-2DPAGE	P54278	HUMAN		
DR	MIM	600259	-		
DR	InterPro	IPR002099	DNA_mis_repair		
DR	InterPro	IPR003594	HATPase_c		
DR	Pfam	PF01118	DNA_mis_repair	1	
DR	Pfam	PF02518	HATPase_c	1	
DR	PROSITE	PS00058	DNA_MISMATCH_REPAIR	1	
KW	DNA_repair	Disease mutation	Anti-oncogene	Polymorphism	
FT	VARIANT	20	R -> O		
FT			/FTID=VAR_004469		
FT	CONFLICT	470	470	P -> S (IN REF.2)	
SO	SEQUENCE	862 AA	95797 MW	B60A605222CBBAC CRC64	

RESULT	15
HS74_CAEL	
ID	HS74_CAEL
AC	P20163
DT	01-FEB-1991 (Rel. 17, Created)
DT	01-FEB-1991 (Rel. 17, Last sequence update)
DT	15-JUL-1998 (Rel. 36, Last annotation update)
DE	Heat shock 70 kDa protein D (Fragment).
GN	HSP-4 OR HSP70D.
OS	Caenorhabditis elegans.
OC	Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae; Poldoderinae; Caenorhabditis.
OX	NCBI_TaxID=6239;

RN
 LN
 RP
 RX
 RA
 RT

[1]
 SEQUENCE FROM N.A.
 MEDLINE=89306577; PubMed=2744444;
 Heschl M.F.P., Baillie D.L.,
 "Identification of a heat-shock pseudogene from *Caenorhabditis*

	Query Match	Best Local Similarity	Score 35.7%	Score 40.0%	DB 1	Length 288
Matches	6	Conservative	6	Mismatches	5	Indels 0; Gaps 0;
QY	3	TFQVEVPGSQHDSQK	19			
	:	:				
Db	133	TFEIDVNGILVSAEDK	149			

Search completed: July 3, 2002, 09:04:43
Job time: 583 sec

Mon Jul 8 07:51:56 2002

us-09-786-648-4.rsp

THIS PAGE BLANK (USPTO)

FT SIGNAL 1 21 POTENTIAL.
FT CHAIN 22 124 CHOLERA TOXIN B PROTEIN (CTB).
SQ SEQUENCE 124 AA; 13919 MW; D6BF83FEF7924EA3 CRC64;

Query Match 100.0%; Score 106; DB 2; Length 124;
Best Local Similarity 100.0%; Pred. No. 2.6e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GATFOVEVPGSHIDSOKKAI 21
DB 66 GATFOVEVPGSHIDSOKKAI 86

RESULT 2

ID O9RP15 PRELIMINARY; PRT; 124 AA.
AC O9RP15; 01-MAY-2000 (TREMblrel. 13, Created)
DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)
DE CHOLERA ENTEROTOXIN B-SUBUNIT.
GN CTXB.
OS Vibrio cholerae.
OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
OX NCBI_TaxId=666;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-KNH002:
RA Shin H.J., Park Y.C., Kim Y.C.;
RT "Cloning and nucleotide sequence analysis of the virulence gene
cassette from Vibrio cholerae KNH002 isolated in Korea."
RL Misalimurhag Hoijl 35:205-210(1999).
DR EMBL; AF175708; AAD51360.1; -
DR HSSP; P01556; 2CHB.
DR Interpro: IPR001835; Enterotoxin_B.
DR Pfam: PF01376; Enterotoxin_B.1.
DR PRINTS: PR00772; ENTEROTOXINB.
DR ProDom: PD012805; Enterotoxin_B.1.
SQ SEQUENCE 124 AA; 13905 MW; 23BF83FEF793E5B9 CRC64;

Query Match 100.0%; Score 106; DB 2; Length 124;
Best Local Similarity 100.0%; Pred. No. 2.6e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GATFOVEVPGSHIDSOKKAI 21
DB 66 GATFOVEVPGSHIDSOKKAI 86

RESULT 3

ID O56635 PRELIMINARY; PRT; 124 AA.
AC O56635; 01-NOV-1996 (TREMblrel. 01, Created)
DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)
DE CHOLERA TOXIN PRECURSOR.
GN CTXB.
OS Vibrio cholerae.
OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
OX NCBI_TaxId=666;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-S7;
RA Yamamoto K., Do V.G., Xu M., Iida T., Miwatani T., Albert M.J.,
RA Honda T.;
RT "Comparison of cholera toxin genes (ctxAB) of non-O1 Vibrio cholerae
strains 854 (O139-bengal) and S7 (O37) from two outbreaks."
RL Submitted (Apr-1994) to the EMBL/GenBank/DBJ databases.
DR EMBL; D30052; BAA06289.1; -
DR HSSP; P01556; 2CHB.

DR Interpro: IPR001835; Enterotoxin_B.
DR Pfam: PF01376; Enterotoxin_B.1.
DR PRINTS: PR00772; ENTEROTOXINB.
DR ProDom: PD012805; Enterotoxin_B.1.
KM signal.
FT SIGNAL 1 21 POTENTIAL.
SQ SEQUENCE 124 AA; 13871 MW; 3F87B2F297953179 CRC64;

Query Match 100.0%; Score 106; DB 2; Length 124;
Best Local Similarity 100.0%; Pred. No. 2.6e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GATFOVEVPGSHIDSOKKAI 21
DB 66 GATFOVEVPGSHIDSOKKAI 86

RESULT 4

ID O93V32 PRELIMINARY; PRT; 124 AA.
AC O93V32; 01-DEC-2001 (TREMblrel. 19, Created)
DT 01-DEC-2001 (TREMblrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
DE HEAT-LABILE ENTEROTOXIN B SUBUNIT.
GN LTH B SUBUNIT.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia
OX NCBI_TaxId=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1032 (ENTEROTOXIGENIC);
RA Komase K.;
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=1032 (ENTEROTOXIGENIC);
RX MEDLINE=95091056; Pubmed=7998417;
RA Tamura S., Asanuma H., Tomita T., Komase K., Kawahara K., Danbara H.,
RA Hattori N., Watanabe K., Suzuki Y., Nagamine T., Aizawa C., Oya A.,
RA Kurata T.;
RT "Escherichia coli heat-labile enterotoxin B subunits supplemented with
RT a trace amount of the holotoxin as an adjuvant for nasal influenza
RT vaccine."
RL Vaccine 12:1083-1089(1994).
DR EMBL; AB011677; BAA25726.1; -
SQ SEQUENCE 124 AA; 14028 MW; 5346BD38B32354C2 CRC64;

Query Match 100.0%; Score 106; DB 2; Length 124;
Best Local Similarity 100.0%; Pred. No. 2.6e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GATFOVEVPGSHIDSOKKAI 21
DB 66 GATFOVEVPGSHIDSOKKAI 86

RESULT 5

ID O94M01 PRELIMINARY; PRT; 124 AA.
AC O94M01; 01-DEC-2001 (TREMblrel. 19, Created)
DT 01-DEC-2001 (TREMblrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
DE CTXB.
GN CTXB.
OS Vibrio phage CTX.
OC Viruses; ssDNA viruses; Inoviridae; Inovirus.
OX NCBI_TaxId=141904;
RN [1]

RP SEQUENCE FROM N.A.
 RA Bhattacharyya T., Nandy R.K., Nair G.B.:
 RT "The entire core region of the ctx-phi (ctx-prophage) in VCE 232, an
 RT environmental strain of *V. cholerae*.";
 RL EMBL: A0414369; to the EMBL/GenBank/DBJ databases.
 DR EMBL: A0414369; AAL09682.1;
 SO SEQUENCE 124 AA; 13884 MW; 8D4C251591B16891 CRC64;

Query Match 100.0%; Score 106; DB 9; Length 124;
 Best Local Similarity 100.0%; Pred. No. 2.6e-10;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GATFOVEVPGSOHIDSOKKAI 21
 Db 66 GATFOVEVPGSOHIDSOKKAI 86

RESULT 6
 ID Q9R646 PRELIMINARY; PRT; 103 AA.
 AC Q9R646;
 DT 01-MAY-2000 (TREMblrel. 13, Created)
 DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)
 DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)
 DE CHOLERA-LIKE ENTEROTOXIN B SUBUNIT.
 OS *Vibrio cholerae*.
 OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; *Vibrio*.
 OX NCBI_TaxID=666;
 RN [1]
 RX MEDLINE=95303036; PubMed=7783690.
 RA Nakashima K., Eguchi Y., Nakasone N.;
 RT "Characterization of an enterotoxin produced by *Vibrio cholerae*
 RT O139.";
 RL Microbiol. Immunol. 39:87-94(1995).
 DR HSSP: P01556; 1XTC.
 DR InterPro: IPR001835; Enterotoxin_B.
 DR Pfam: PF01376; Enterotoxin_B.
 DR PRINTS: PR00772; ENTEROTOXIN_B.
 DR PRODOM: PD012805; Enterotoxin_B; 1.
 SO SEQUENCE 103 AA; 11645 MW; 992A05C077F2F70E CRC64;

Query Match 94.3%; Score 100; DB 2; Length 103;
 Best Local Similarity 95.2%; Pred. No. 2.2e-09;
 Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GATFOVEVPGSOHIDSOKKAI 21
 Db 45 GATFOVEVPGSOHIDSOKKAI 65

RESULT 7
 ID Q9LJB0 PRELIMINARY; PRT; 552 AA.
 AC Q9LJB0;
 DT 01-OCT-2000 (TREMblrel. 15, Created)
 DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)
 DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
 DE EMBL/CAB7996.1.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-COLUMBIA.
 RA Kaneo T., Kato T., Sato S., Nakamura Y., Asamizu E., Tabata S.;
 RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
 SO SEQUENCE FROM N.A.

RC STRAIN-COLUMBIA;
 RX MEDLINE=20363099; PubMed=10907853;
 RA Nakamura Y.;
 RT "Structural analysis of Arabidopsis thaliana chromosome 3. II.
 RT Sequence features of the regions of 4,251,695 bp covered by ninety pl,
 RT TAC and BAC clones.";
 RL DNA Res. 7:217-221(2000).
 DR EMBL: AP000606; BAB01195.1;
 SO SEQUENCE 552 AA; 63036 MW; 740EA16CDBB2447 CRC64;

Query Match 49.1%; Score 52; DB 10; Length 552;
 Best Local Similarity 58.8%; Pred. No. 2.3;
 Matches 10; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 5 QVEVPGSOHIDSOKKAI 21
 Db 147 QVEVPGSOHIDSOKKAI 163

RESULT 8
 ID Q99987 PRELIMINARY; PRT; 508 AA.
 AC Q99987;
 DT 01-MAY-1997 (TREMblrel. 03, Created)
 DT 01-JAN-1998 (TREMblrel. 05, Last sequence update)
 DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
 DE VRK2.
 OS *Homo sapiens* (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=LIVER;
 RX MEDLINE=98008921; PubMed=9344656;
 RA Nezu J., Oku A., Jones M.H., Shimane M.;
 RT "Identification of two novel human putative serine/threonine kinases,
 RT VRK1 and VR K2, with structural similarity to Vaccinia virus BIR
 RT kinase.";
 RL Genomics 45:327-331(1997).
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 DR EMBL: AB000450; BAA19109.1;
 DR HSSP: Q06486; ICKI.
 DR InterPro: IPR000719; Euk_pkinase.
 DR InterPro: IPR002290; Ser_thr_pkinase.
 DR Pfam: PF00069; pkinase.2.
 DR PROSITE: PS5011; PROTEIN_KINASE_DOM.1.
 DR PROSITE: PS00108; PROTEIN_KINASE_ST.1.
 DR ATP-binding; Serine/threonine-protein kinase; Transferase.
 SO SEQUENCE 508 AA; 58126 MW; 157FBF8F48511A4 CRC64;

Query Match 48.1%; Score 51; DB 4; Length 508;
 Best Local Similarity 50.0%; Pred. No. 3.1;
 Matches 10; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

Qy 1 GATFOVEVPGSOHIDSOKKAI 20
 Db 329 GQSTVWHTPNSOKYDSOKKA 348

RESULT 9
 ID Q9LGM2 PRELIMINARY; PRT; 428 AA.
 AC Q9LGM2;
 DT 01-OCT-2000 (TREMblrel. 15, Created)
 DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)
 DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
 DE ESTS A0056822 (S20908).
 OS *Oryza sativa* (Rice).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoideae; Oryzae; Oryza.
 OX NCBI_TaxID=4530;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CV. NIPPONBARE;
 RA Sasaki T., Matsumoto T., Yamamoto K.;
 RT "Oryza sativa nippobare(GA3) genomic DNA, chromosome 1, PAC
 clone: P0041E11.";
 RT Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CV. NIPPONBARE;
 RA Sasaki T., Matsumoto T., Yamamoto K.;
 RT "Oryza sativa nippobare(GA3) genomic DNA, chromosome 1, PAC
 clone: P0433F09.";
 RT Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AP002521; BAA96762.1; -;
 DR EMBL: AP002539; BAB08201.1; -;
 DR HSSP: Q06319; 1BUC.
 DR InterPro: IPR001552; Acyl-CoA_dh.
 DR Pfam: PF00441; Acyl-CoA_dh.1.
 DR Pfam: PF02770; Acyl-CoA_dh.M.1.
 DR Pfam: PF02771; Acyl-CoA_dh.N.1.
 DR PROSITE: PS00073; ACYL_COA_DH_2; UNKNOWN.1.
 DR PROSITE: PS00073; ACYL_COA_DH_2; UNKNOWN.1.
 SQ SEQUENCE 428 AA; 46132 MW; 8D34E3698A8E6367 CRC64;

Query Match 46.2%; Score 49; DB 10; Length 428;
 Best Local Similarity 52.6%; Pred. No. 5.6;
 Matches 10; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

OY 3 TFEVPEVPGSQHIDSOKKAI 21
 Db 183 TTTATKTVGGWHIDGQKRWI 201

RESULT 10
 ID Q40742 PRELIMINARY; PRT; 392 AA.
 AC Q40742;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE OSRAD23.
 OS Oryza sativa (Rice).
 OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoideae; Oryzae; Oryza.
 OX NCBI_TaxID=4530;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-NIPPONBARE;
 RX MEDLINE=97369378; PubMed=9225866;
 RA Schultz T.F., Quattrano R.S.;
 RT "Characterization and expression of a rice RAD23 gene.";
 RT Plant Mol. Biol. 34:557-562(1997).
 DR EMBL: U63530; AAB65841.1; -;
 DR HSSP: P54725; 1DVO.
 DR InterPro: IPR000449; UBA.
 DR InterPro: IPR000626; Ubiquitin.
 DR Pfam: PF00627; UBA.2.
 DR Pfam: PF00240; ubiquitin.1.
 DR SMART: SM00165; UBA.2.
 DR SMART: SM00213; UBO.1.
 DR PROSITE: PS00053; UBQUITIN_2; 1.
 SQ SEQUENCE 392 AA; 41754 MW; BDE08574CC7CACB CRC64;

Query Match 45.3%; Score 48; DB 10; Length 392;
 Best Local Similarity 42.9%; Pred. No. 7.6;
 Matches 9; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

OY 1 GATFVEVPGSQHIDSOKKAI 21
 Db 10 GSTFQIEVDSAKYADVKKRII 30

RESULT 11
 ID Q9NKD5 PRELIMINARY; PRT; 395 AA.
 AC Q9NKD5;
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
 DE HYPOTHETICAL 43.6 KDA PROTEIN.
 GN BG:DS01514.3.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-Y, CN BW SP;
 RX MEDLINE=99403001; PubMed=10471707;
 RA Ashburner M., Mitra S., Roote J., Lewis S.E., Blazej R., Davis T.,
 RA Doyle C., Galle R., George R., Harris N., Hartzell G., Harvey D.,
 RA Hong L., Houston K., Hoskins R., Johnson G., Martin C., Moshrefi A.,
 RA Palazzolo M., Reese M.G., Spradling A., Tsang G., Wan K., Whitelaw K.,
 RA Celisner S., Rubin G.M.;
 RT "An exploration of the sequence of a 2.9-kb region of the genome of
 Drosophila melanogaster: the Adh region.";
 RT Genetics 153:179-219(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-Y, CN BW SP;
 RA Celisner S.E., Agbayani A., Arcaine T.T., Baxter E., Blazej R.G.,
 RA Butenoff C., Champe M., Chavez C., Chew M., Ciesiolka L., Doyle C.M.,
 RA Farfan D.E., Galle R., George R.A., Harris N.L., Hoskins R.A.,
 RA Houston K.A., Humastil S.R., Kaira K., Kearney L., Kim E., Lee B.,
 RA Lewis S., Li P., Lomston M.A., Mazda P., Moshrefi A.R., Moshrefi M.,
 RA Nixon K., Pacleb J.M., Park S., Pfeiffer B., Poon L., Sequeira A.,
 RA Sethi H., Snir E., Swirskas R.R., Wan K.H., Weinburg T., Zhang R.,
 RA Zieran L.L., Rubin G.M.;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AB003408; AAF44847.1; -;
 DR Flybase: FBgn0028907; BG:DS01514.3.
 KW Hypothetical protein.
 SQ SEQUENCE 395 AA; 43561 MW; AEF1CC4ADD3DA73 CRC64;

Query Match 44.3%; Score 47; DB 5; Length 395;
 Best Local Similarity 56.2%; Pred. No. 11;
 Matches 9; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

OY 5 QVEVPGSQHIDSOKKA 20
 Db 329 RVSPGSTHIDADANA 344

RESULT 12
 ID Q9PAE0 PRELIMINARY; PRT; 1166 AA.
 AC Q9PAE0;
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE PHEROMONE-RESPONSIVE MAPKK KINASE UB4.
 GN UB4.
 OS Ustilago maydis (Smut fungus).
 OC Eukaryota; Fungi; Basidiomycota; Ustilaginomycetes;
 OC Ustilaginomycetidae; Ustilaginales; Ustilaginaceae; Ustilago.
 OX NCBI_TaxID=5270;
 RN [1]
 RP SEQUENCE FROM N.A.

RC STRAIN-1/2;
 RX MEDLINE-20331594; PubMed-10875339;
 RA Andrews D.L., Egan J.D., Mayorga M.E., Gold S.E.;
 RT "The Ustilago maydis ubc4 and ubc5 genes encode members of a MAP
 RT kinase cascade required for filamentous growth."
 RL Mol. Plant Microbe Interact. 13:781-786(2000).
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 DR EMBL: AF197562; AAF86841.1;
 DR HSSP: P24941; 1B38.
 DR InterPro: IPR000719; Euk_kinase.
 DR InterPro: IPR000159; RA.
 DR InterPro: IPR002290; Ser_thr_kinase.
 DR Pfam: PF00069; pkinase; 1.
 DR SMART: SM00220; S_TKc; 1.
 DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
 DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
 KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
 SQ SEQUENCE 1166 AA; 123298 MW; A89E0019AABD4172 CRC64;

Query Match 43.4%; Score 46.5; DB 3; Length 1166;
 Best Local Similarity 47.6%; Pred. No. 47;
 Matches 10; Conservative 4; Mismatches 6; Indels 1; Gaps 1;
 OY 2 ATFOVEVPGSOHIDSOKKAI 21
 DB 729 AKQVELPSGDSHDOCKKM 749

RESULT 13
 ID 049876 PRELIMINARY; PRT; 293 AA.
 AC 049876;
 DT 01-JUN-1998 (TREMblrel. 06, Created)
 DT 01-JUN-1998 (TREMblrel. 06, Last sequence update)
 DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
 DE CLASS III CHITINASE PRECURSOR (EC 3.2.1.14).
 GN LUPAL.
 OS Lupinus albus (White lupine).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids I; Fabales; Fabaceae; Papilionoideae; Genisteae; Lupinus.
 OX NCBI_TaxID=3870;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CV. ULTRA;
 RA Regalado A.P., Vidal S., Neves A., Ricardo C.P.P.,
 RA Rodrigues-Pousada C.;
 RT "Constitutive and stress-induced expression of a class III chitinase
 RT from Lupinus albus."
 RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: BELONGS TO CHITINASE CLASS II (FAMILY 18 OF GLYCOSYL
 CC HYDROLASES).
 DR EMBL: Y16415; CAA76203.1;
 DR HSSP: P23472; 2HVM.
 DR InterPro: IPR001579; Chitinase_2.
 DR Pfam: PF00192; chitinase_2; 1.
 DR PROSITE: PS01095; CHITINASE_18; 1.
 KW Glycosidase; Hydrolase; Signal.
 FT SIGNAL 1 27
 FT CHAIN 28 293
 FT SEQUENCE 293 AA; 31128 MW; DB1B21728F657F2F CRC64;

Query Match 43.4%; Score 46; DB 10; Length 293;
 Best Local Similarity 38.1%; Pred. No. 12;
 Matches 8; Conservative 4; Mismatches 9; Indels 0; Gaps 0;
 OY 1 GATFOVEVPGSOHIDSOKKAI 21
 DB 142 GIDFIEAGAGOHYDELARAL 162

RESULT 14
 ID 0903M9 PRELIMINARY; PRT; 374 AA.
 AC 0903M9;
 DT 01-MAY-2000 (TREMblrel. 13, Created)
 DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)
 DT 01-MAR-2001 (TREMblrel. 16, Last annotation update)
 DE C40H5.3 PROTEIN.
 GN C40H5.3.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidae;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA White S.;
 RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-99069613; PubMed-9851916;
 RA none;
 RT "Genome sequence of the nematode C.elegans: A platform for
 RT investigating biology."
 RL Science 282:2012-2018(1998).
 DR EMBL: Z81482; CAB03954.2;
 SQ SEQUENCE 374 AA; 42198 MW; 67D202886DA7824 CRC64;

Query Match 43.4%; Score 46; DB 5; Length 374;
 Best Local Similarity 42.1%; Pred. No. 16;
 Matches 8; Conservative 2; Mismatches 9; Indels 0; Gaps 0;
 OY 1 GATFOVEVPGSOHIDSOKK 19
 DB 230 GVEFHWFPGOPHAEQKE 248

RESULT 15
 ID 095XK9 PRELIMINARY; PRT; 574 AA.
 AC 095XK9;
 DT 01-MAY-2000 (TREMblrel. 13, Created)
 DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)
 DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
 DE T17H3.2 PROTEIN (AT1G27520/T17H3_2).
 GN T17H3.2.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CV. COLUMBIA;
 RA Vysotskaia V.S., Schwartz J.R., Yu G., Toriumi M., Lenz C., Liu S.,
 RA Lee J.M., Liu A., Li J., Gonzalez A., Liu K., Vaysberg M., Sakano H.,
 RA Chin C., Howng B., Choi E., Chlou J., Altafi H., Araujo R., Brooks S.,
 RA Buehler E., Chao Q., Conn L., Conway A.B., Dunn P., Hansen N.,
 RA Huizer L., Khan S., Kim C., Palm C., Rowley D., Shin P., Walker M.,
 RA Davis R.W., Ecker J.R., Federspiel N.A., Theologis A.;
 RT "Arabidopsis thaliana chromosome 1 BAC T17H3 sequence."
 RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Cheuk R., Chen H., Kim C.J., Koesema E., Meyers M.C., Banh J.,
 RA Bowser L., Carinini P., Dale J.M., Goldsmith A.D., Hayashizaki Y.,
 RA Ishida J., Jiang P.X., Jones T., Kamiya A., Karlin-Neumann G.,
 RA Kawai J., Lam B., Lee J.M., Lin J., Liu S.X., Miranda M., Natrusaka M.,
 RA Nguyen M., Onodera C.S., Palm C.J., Pham P.K., Quach H.L., Sakurai T.,
 RA Satou M., Seki M., Southwick A., Tang C.C., Toriumi M., Yamada K.,
 RA Yamamura Y., Yu G., Yu S., Shinozaki K., Davis R.W., Theologis A.,
 RA Ecker J.R.;

	12	107	100.0	134	22	AAB73241	Recombinant exotoxin
	13	107	100.0	142	22	AAB73242	Recombinant exotoxin
	14	107	100.0	155	22	AAB73243	Recombinant exotoxin
	15	107	100.0	163	22	AAB73244	Recombinant exotoxin
	16	107	100.0	371	20	AAV01300	Labile toxin (LT-B)
	17	107	100.0	371	20	AAW67443	C. jejuni flagellin
	18	101	94.4	21	21	AAV87462	Cholera toxin B/en
	19	101	94.4	103	17	AAW04857	Synthetic cholera
	20	101	94.4	103	17	AAK94939	Heat labile entero
	21	101	94.4	103	17	AAW06606	Cholera toxin B su
	22	101	94.4	103	17	AAW06607	Cholera toxin B su
	23	101	94.4	103	19	AAW08008	Amino acid sequenc
	24	101	94.4	103	22	AAB62367	V. cholera cholerat
	25	101	94.4	103	22	AAB62370	V. cholera cholerat
	26	101	94.4	105	22	AAB62369	V. cholera cholerat
	27	101	94.4	118	11	AAAR04163	Cholera Toxin B-su
	28	101	94.4	123	11	AAAB62374	E. coli LTB protei
	29	101	94.4	124	10	AAAP93561	B subunit of the h
	30	101	94.4	124	17	AAW06605	Cholera toxin B su
	31	101	94.4	124	19	AAW59770	Amino acid sequenc
	32	101	94.4	124	21	AAV96652	Plant-optimized E.
	33	101	94.4	124	22	AAAG65992	Cholera toxin B su
	34	101	94.4	124	22	AAAB62359	V. cholera strain
	35	101	94.4	124	22	AAAB62363	V. cholera cholerat
	36	101	94.4	124	22	AAAB62368	V. cholera cholerat
	37	101	94.4	124	22	AAAB62376	E. coli LTB protei
	38	101	94.4	124	22	AAAB62378	E. coli LTB protei
	39	101	94.4	125	22	AAAB62372	E. coli LTB protei
	40	101	94.4	131	11	AAAR04825	LTB-CTB fusion pro
	41	101	94.4	138	15	AAAS0237	Sequence of LT-B-N
	42	101	94.4	170	20	AAW94082	LTB-CTP fusion pro
	43	101	94.4	360	22	AAU00507	E. coli heat-labili
	44	101	94.4	382	22	AAU00506	E. coli heat-labili
	45	101	94.4	405	12	AAAR1272	HSV-1 antigen/heat

ALIGNMENTS

RESULT	1	
AA87463		
ID	AA87463 standard; peptide; 21 AA.	
AC	AA87463;	
XX		
DT	03-JUL-2000 (first entry)	
DE	E. coli heat labile enterotoxin B-derived peptide, SEQ ID NO:5.	
XX		
XX	Heat labile enterotoxin subunit B; EtXB;	
KW	beta-4-alpha-2 loop; GM-1 ganglioside receptor; immunomodulation;	
KW	adjuvant; immune disorder; diarrhoea.	
XX		
XS	Escherichia coli.	
NN		
PN	WO200014114-A1.	
XX		
PD	16-MAR-2000.	
XX		
PE	07-SEP-1999; 99WO-GH02970.	
XX		
PR	07-SEP-1998; 98GB-0019484.	
XX		
PA	(UYBR-) UNIV BRISTOL.	
XX		
PI	Williams NA, Hirst TR;	
DR	WPI; 2000-256943/22.	
PT	Derivatives of Escherichia coli heat labile enterotoxins useful as	
FT	immunomodulators and for treating diarrhea and which do not bind the	
PT	glycolipid receptor GM-1 -	
XX		

PS Disclosure; Page 15; 62pp; English.
 CC The invention relates to peptide fragments of the Escherichia coli heat
 CC labile enterotoxin (Etx) and its closely related homologue, cholera
 CC toxin (Ctx) from Vibrio cholerae which do not bind to the ubiquitous
 CC GM-1 ganglioside receptors on host cell surfaces. Both Etx and Ctx are
 CC composed of one A subunit and five identical B subunits. The A subunit
 CC is responsible for toxicity, possessing adenosine diphosphate (ADP) ADP-
 CC ribosyltransferase activity, while the B subunits (EtxB and CtxB)
 CC facilitate the entry of subunit A into the host cell via the binding and
 CC cross-linking of GM-1 receptors. Although GM-1 binding is responsible
 CC for some of the effects of Etx and Ctx, it has been found that certain
 CC effects of the toxins, such as immunomodulation, are not mediated
 CC through GM-1 binding. The peptides of the invention are fragments of the
 CC beta-4- alpha-2 loop of EtxB and/or CtxB, exerting the same effects as
 CC normal EtxB and CtxB subunits, except that they do not bind or cross link
 CC GM-1. They may be used in medicine as an immunomodulator or adjuvant.
 CC They may also be used as an inhibitor for toxin-induced diarrhoea.
 CC Therefore, the peptides may be used in the production of a composition
 CC for treating, preventing and/or modulating a disease associated with an
 CC immune disorder and/or toxin-induced diarrhoea. Sequences AAY87460-Y87463
 CC represent preferred peptides of the invention, AAY87460 being
 CC particularly preferred.
 CC
 SQ Sequence 21 AA;

Query Match 100.0%; Score 107; DB 21; Length 21;
 Best Local Similarity 100.0%; Pred. No. 2.6e-11;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GEFQVEVPSQHIDSQKAI 21
 |||||
 DB 1 getfqvevpsqhidsqkai 21

RESULT 2

AAR72545
 ID AAR72545 standard; peptide; 93 AA.

XX
 AC AAR72545;

XX
 DT 28-NOV-1995 (first entry)

DE ADP-ribosylating toxin (verotoxin-1 B-subunit).

XX
 KM ADP-ribosylating toxin; pertussis holotoxin; B-subunit;

KW active site; E. coli heat labile toxin; verotoxin-1;

KM Bordetella pertussis vaccines.

XX
 OS Bacteria sp.

XX
 PN EP646599-A.

XX
 PD 05-APR-1995.

XX
 PF 23-AUG-1994; 94EP-0306219.

XX
 PR 24-AUG-1993; 93US-0110947.

XX
 PR 31-MAY-1994; 94US-0251121.

XX
 PA (CONN-) CONNUGHT LAB LTD.

XX
 PA (UYAL-) UNIV ALBERTA.

XX
 PI Armstrong GD, Cockle SA, Hazes B, Klein MH, Loosmore S;

XX
 PI Oomen R, Read RJ, Stein PE;

XX
 DR WPI; 1995-132623/18.

XX
 PR New modified forms of pertussis holotoxin - developed using

XX
 PR crystalline forms of pertussis holotoxin and its complexes with

XX
 PR other molecules

PS Disclosure; Fig 5; 54pp; English.
 CC AAR72540-R72545 are structurally equivalent B-subunits from three
 CC ADP-ribosylating toxins, pertussis holotoxin (PT), E. coli heat
 CC labile toxin (LT), and verotoxin-1 (VT). The structural
 CC information obtained from these comparisons was used to identify
 CC sites which contribute to PT's biological activity. By modifying
 CC these sites the claimed PT mutants of the invention were produced,
 CC they can be used in the development of vaccines against Bordetella
 CC pertussis infection.
 CC
 SQ Sequence 93 AA;

Query Match 100.0%; Score 107; DB 16; Length 93;
 Best Local Similarity 100.0%; Pred. No. 1.5e-10;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GEFQVEVPSQHIDSQKAI 21
 |||||
 DB 35 getfqvevpsqhidsqkai 55

RESULT 3

AAY1816
 ID AAY1816 standard; peptide; 93 AA.

XX
 AC AAY1816;

XX
 DT 08-DEC-1999 (first entry)

DE Escherichia coli verotoxin-1 B-subunit.

XX
 KM ADP-ribosylating toxin; PT; crystalline pertussis holotoxin; ETA; DT;

KW three-dimensional structure; LT; immunoprotective; infection.

XX
 OS Escherichia coli.

XX
 PN US5965385-A.

XX
 PD 12-OCT-1999.

XX
 PF 06-JUN-1995; 95US-0467974.

XX
 PR 22-AUG-1994; 94US-0292968.

XX
 PR 24-AUG-1993; 93US-0110947.

XX
 PR 31-MAY-1994; 94US-0251121.

XX
 PA (CONN-) CONNUGHT LAB LTD.

XX
 PA (UYAL-) UNIV ALBERTA.

XX
 PI Read RJ, Stein PE, Cockle SA, Loosmore S, Klein MH, Armstrong GD;

XX
 PI Hazes B, Oomen RP;

XX
 DR WPI; 1999-579908/49.

XX
 DR WPI; 1999-579908/49.

XX
 DR WPI; 1999-579908/49.

XX
 DR WPI; 1999-579908/49.

XX
 DR WPI; 1999-579908/49.

XX
 DR WPI; 1999-579908/49.

XX
 DR WPI; 1999-579908/49.

XX
 DR WPI; 1999-579908/49.

XX
 DR WPI; 1999-579908/49.

Example 3; Fig 5; 41pp; English.

New method for producing modified pertussis holotoxin -

A method has been developed of producing a modified pertussis holotoxin,
 involving analysis of the 3-dimensional form of the crystalline
 holotoxin. The pertussis holotoxin modification process comprises:
 (1) identification of at least one amino acid (aa) residue of the
 holotoxin for modification by analysing the 3-dimensional form of the
 crystalline holotoxin, in relation to known information of the protein
 structure and function; (2) effecting mutagenesis (by removing or
 replacing a nucleotide sequence encoding at least one (aa)) of a tox
 operon; and (3) expressing mutant tox box in a Bordetella organism to
 produce the modified holotoxin. This method is used for modifying
 pertussis holotoxin, by studying its 3-dimensional crystalline
 structure. Modifying the holotoxin, alters its biological properties.
 By analysing the 3-dimensional crystalline structure of the pertussis

CC holotoxin, functional (aa) which affect biological properties of the
 CC pertussis holotoxin can be identified. This can be used to predict (aa)
 CC which contribute to the toxicity of the holotoxin to produce
 CC immunoprotective, genetically-detoxified analogues of pertussis
 CC holotoxin. The present sequence represents an ADP-ribosylating toxin
 CC B-subunit peptide used in the exemplification of the present
 CC invention.

XX Sequence 93 AA;

Query Match 100.0%; Score 107; DB 20; Length 93;
 Best Local Similarity 100.0%; Pred. No. 1.5e-10;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GETFOVEVPGSOHIDSOKKAI 21
 |||||
 Db 35 getfqvpgsqhidsqkai 55

RESULT 4

AAW95226 standard; peptide; 93 AA.

XX AAW95226;
 AC
 XX 16-MAR-1999 (first entry)
 DT
 XX E. coli heat-labile toxin (LT) beta-subunit sequence.
 DE
 XX Pertussis holotoxin; PT: modified; effector; toxicity; cell binding;
 KM enzyme; ADP-ribosylation; toxin; crystalline; X-ray crystallography;
 KW structural analysis; interacting site; mitogenicity; adjuvantility;
 KM heat-labile; LT.
 KW
 XX Escherichia coli.
 OS
 XX US856122-A.
 PN
 XX 05-JAN-1999.
 PD
 XX 22-AUG-1994; 94US-0292968.
 PF
 XX 22-AUG-1994; 94US-0292968.
 PR 24-AUG-1993; 93US-0110947.
 PR 31-MAY-1994; 94US-0251121.
 XX
 XX (UNIV ALBERTA).

PI Armstrong GD, Cockle SA, Hazes B, Klein MH, Loosmore S;
 PI Oomen RP, Read RJ, Stein PE;
 PI
 XX
 DR WPI: 1999-105104/09.

PT Modifications to e.g. enzymatic activity, mitogenicity and cell
 binding of pertussis holotoxin - by identifying interaction sites of
 PT a molecule with crystalline toxin and modifying the identified site
 XX
 PS Example 3; Fig 5; 40pp; English.

XX The invention relates to methods of preparing a pertussis holotoxin (PT)
 CC having a modified biological activity. One method comprises identifying
 CC at least 1 site in a PT that interacts with a molecule that is capable of
 CC forming a complex with the holotoxin and which molecule is an effector
 CC molecule which is an adenine nucleotide and which site contributes to
 CC toxicity, cell binding or enzymatic activity of PT. The functional
 CC interacting site(s) are identified by analysing the three dimensional
 CC structure of crystalline PT, determined by X-ray crystallography. The
 CC identified interacting site(s) are modified to alter toxicity, cell
 CC binding or enzyme activity of the PT. The methods can be used to alter a
 CC biological activity such as toxicity, enzymatic activity, mitogenicity,
 CC cell binding and adjuvantility of the PT. The three-dimensional structure
 CC of PT have functional and/or structural resemblance to other bacterial

CC toxins such as diphtheria toxin (DT), Pseudomonas exotoxin A (ETA), the
 CC heat-labile toxin of E. coli (LT) and verotoxin-1 (VT). The present
 CC sequence represents the beta-subunit of LT toxin.

XX Sequence 93 AA;

Query Match 100.0%; Score 107; DB 20; Length 93;
 Best Local Similarity 100.0%; Pred. No. 1.5e-10;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GETFOVEVPGSOHIDSOKKAI 21
 |||||
 Db 35 getfqvpgsqhidsqkai 55

RESULT 5

AAV68365 standard; peptide; 93 AA.

XX AAV68365;
 AC
 XX 17-APR-2000 (first entry)
 DT
 XX Heat labile toxin B subunit SEQ ID NO:26.
 DE
 XX Pertussis toxin; heat labile toxin; verotoxin; exotoxin A; holotoxin;
 KW diphtheria toxin; ADP-ribosylating toxin; mannose binding protein;
 KW infection; crystal structure; X-ray crystallography; detoxification;
 KW immunogenic.
 KW
 XX Escherichia coli.
 OS
 XX US6018022-A.
 PN
 XX 25-JAN-2000.
 PD
 XX 06-JUN-1995; 95US-0467976.
 PF
 XX 22-AUG-1994; 94US-0292968.
 PR 24-AUG-1993; 93US-0110947.
 PR 31-MAY-1994; 94US-0251121.
 XX
 XX (CONNAUGHT LAB LTD.
 PA (UNIV ALBERTA).

PI Read RJ, Stein PE, Cockle SA, Loosmore S, Klein MH, Armstrong GD;
 PI Hazes B, Oomen RP;
 PI
 XX
 DR WPI: 2000-136703/12.

PT Crystalline form of isolated pertussis holotoxin useful in studying
 PT proteins which have functional resemblance -
 XX
 PS Example 3; Fig 5; 42pp; English.

XX The present invention describes a crystalline form of isolated
 CC pertussis holotoxin, in which the molecules of pertussis toxin have
 CC a three dimensional structure represented in the specification,
 CC complexed with a polysaccharide molecule capable of forming a complex
 CC with the holotoxin. The crystalline form of the pertussis holotoxin
 CC can be used in a comparison with other proteins which have functional
 CC resemblance to pertussis holotoxin with the aim of modifying other
 CC proteins. Identifying the unknown sites of toxicity by comparison
 CC with the three dimensional structure of pertussis holotoxin provides a
 CC technique for detoxification of toxins to produce useful immunogenic
 CC but non-toxic analogues. It can also be used as a primary standard to
 CC measure the quantity, purity or efficacy of less pure compositions
 CC containing pertussis toxin. AAV68340 to AAV68385 represent peptides
 CC used in the exemplification of the present invention.

XX Sequence 93 AA;

Query Match 100.0%; Score 107; DB 21; Length 93;
Best Local Similarity 100.0%; Pred. No. 1.5e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GETFEVPEVPGSHIDSQKAI 21
DB 35 getfgevpgshidsqkai 55

RESULT 6
AAB6239
ID AAB6239 standard; Protein; 93 AA.

AC AAB6239;
DT 03-APR-2001 (first entry)
DE E coli verotoxin-1 B subunit SEQ ID NO: 26.

XX Pertussis toxin; crystal structure; whooping cough; biological activity;
KM lymphocytosis-promoting factor; histamine-sensitising factor;
KM islet-activating protein.

XX Escherichia coli.

XX US6168928-B1.

XX 02-JAN-2001.

XX 21-MAY-1998; 98US-0082514.

XX 22-AUG-1994; 94US-0292968.

XX 24-AUG-1993; 93US-0110947.

XX 31-MAY-1994; 94US-0251121.

XX (CONN-) CONNACHT LAB LTD.

XX Read RJ, Cockle SA, Oomen RP, Loosmore S, Klein MH, Armstrong GD;
PI Hazen B, Stein PE;

XX WPI; 2001-122260/13.

XX Modifying pertussis holotoxin to produce detoxified PT analogs,
PT comprising analyzing crystalline structure of toxin, to identify sites
of toxicity, cell binding or enzyme activity of PT and modifying
PT identified site -

XX Example 3; Fig 5; 41pp; English.

XX The present invention provides a method for producing a pertussis toxin
CC (also designated lymphocytosis-promoting factor, histamine-sensitising
CC factor and islet activating protein) with a modified biological activity,
CC involving analysing the crystal structure of the protein to identify
CC active sites which can then be modified. This may lead to an alteration
CC in the toxicity, cell binding or enzyme activity of the toxin. This can
CC be used in the production of immunoprotective analogues of pertussis
CC toxin. Pertussis toxin is the cause of whooping cough following infection
CC by Bordetella pertussis.

XX Sequence 93 AA;

Query Match 100.0%; Score 107; DB 22; Length 93;
Best Local Similarity 100.0%; Pred. No. 1.5e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GETFEVPEVPGSHIDSQKAI 21
DB 35 getfgevpgshidsqkai 55

RESULT 7

AAB62379
ID AAB62379 standard; Protein; 103 AA.

XX AAB62379;

DT 29-JUN-2001 (first entry)

XX E. coli LTB protein variant (GI: 494265).

XX ABB5B subunit protein; mutation; recombinant; cholera toxin B; CTB;
KM immunogen; autoimmune disease; neuron demyelinating disease; rCTB; LTB;
KM infection; parasitic; immunosuppressive; antiarthritic; antineumatic;
KM antidiabetic; neuroprotective; vaccine; heat labile toxin B; variant.

XX Escherichia coli.

XX WO200127144-A2.

XX 19-APR-2001.

XX 05-OCT-2000; 2000WO-US27607.

XX 08-OCT-1999; 99US-0158561.

XX (ACTI-) ACTIVE BIOTECH AB.

XX Handley HH, Haaparanta T, Ewalt KL;

XX WPI; 2001-281974/29.

XX Recombinant ABB5B subunit protein comprising a mutation that alters the
PT number residues available for chemical modification, useful for
PT covalently linking to an immunogen or vaccine which can be used for
PT treating autoimmune diseases -

XX Disclosure; Page 72-73; 78pp; English.

XX The invention relates to a recombinant ABB5B subunit protein comprising at
CC least one mutation, where the mutation alters the number of amino acid
CC residues available for chemical modification as compared to a wild type

CC ABB5B subunit protein, and where the recombinant protein retains an
CC effective target ligand binding affinity. A recombinant ABB5B subunit
CC protein such as cholera toxin B protein (CTB) can be specifically
CC covalently linked at lysines or cysteines to an immunogen or vaccine.
CC Recombinant mutant CTB (rCTB) can be used in the treatment of autoimmune
CC diseases e.g. rheumatoid arthritis, encephalomyelitis (or other neuron
CC demyelinating diseases) and diabetes. The rCTB or other B subunits of the
CC invention can also be used to induce tolerance to infection, e.g.
CC parasitic infection. The present sequence represents a variant of the
CC E. coli heat labile toxin B (LTB) protein.

XX Sequence 103 AA;

Query Match 100.0%; Score 107; DB 22; Length 103;
Best Local Similarity 100.0%; Pred. No. 1.7e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GETFEVPEVPGSHIDSQKAI 21
DB 45 getfgevpgshidsqkai 65

RESULT 8
AAB62377
ID AAB62377 standard; Protein; 123 AA.

XX AAB62377;

DT 29-JUN-2001 (first entry)

XX E. coli LTB protein variant (GI: 223254).

KM AB5B subunit protein; mutation; recombinant; cholera toxin B; CTB;
KM immunogen; autoimmune disease; neuron demyelinating disease; rCTB; LTB;
KM infection; parasitic; immunosuppressive; antiarthritic; antineumatic;
KM antidiabetic; neuroprotective; vaccine; heat labile toxin B; variant.
OS Escherichia coli.
XX
XX WO200127144-A2.
XX
XX 19-APR-2001.
XX
XX 05-OCT-2000; 2000WO-US27607.
XX
XX 08-OCT-1999; 99US-0158561.
XX
XX (ACTI-) ACTIVE BIOTECH AB.
XX
XX Handley HH, Haaparanta T, Ewalt KL;
XX
XX WPI; 2001-281974/29.
XX
XX
XX Recombinant AB5B subunit protein comprising a mutation that alters the
PT number residues available for chemical modification, useful for
PT covalently linking to an immunogen or vaccine which can be used for
PT treating autoimmune diseases -
XX
XX
XX Disclosure; Page 72; 78pp; English.
XX
XX The invention relates to a recombinant AB5B subunit protein comprising at
CC least one mutation, where the mutation alters the number of amino acid
CC residues available for chemical modification as compared to a wild type
CC AB5B subunit protein, and where the recombinant protein retains an
CC effective target ligand binding affinity. A recombinant AB5B subunit
CC protein such as cholera toxin B protein (CTB) can be specifically
CC covalently linked at lysines or cysteines to an immunogen or vaccine.
CC Recombinant mutant CTB (rCTB) can be used in the treatment of autoimmune
CC diseases e.g. rheumatoid arthritis, encephalomyelitis (or other neuron
CC demyelinating diseases) and diabetes. The rCTB or other B subunits of the
CC invention can also be used to induce tolerance to infection, e.g.
CC parasitic infection. The present sequence represents a variant of the
CC E. coli heat labile toxin B (LTB) protein.
XX
XX
XX Sequence 123 AA;
SQ

Query Match 100.0%; Score 107; DB 22; Length 123;
Best Local Similarity 100.0%; Pred. No. 2.1e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GETFQVEVPSQHIDSOKKAI 21
Db 65 getfqvevpsqghidsqkai 85
|||||

RESULT 9
AAB62373 standard; Protein; 124 AA.
XX
XX AAB62373;
XX
XX 29-JUN-2001 (first entry)
XX
XX E. coli LTB protein variant (GI: 1169505).
XX
XX
XX AB5B subunit protein; mutation; recombinant; cholera toxin B; CTB;
KM immunogen; autoimmune disease; neuron demyelinating disease; rCTB; LTB;
KM infection; parasitic; immunosuppressive; antiarthritic; antineumatic;
KM antidiabetic; neuroprotective; vaccine; heat labile toxin B; variant.
XX
XX Escherichia coli.
XX
XX WO200127144-A2.
XX
XX

PD 19-APR-2001.
XX
XX 05-OCT-2000; 2000WO-US27607.
XX
XX 08-OCT-1999; 99US-0158561.
XX
XX (ACTI-) ACTIVE BIOTECH AB.
XX
XX Handley HH, Haaparanta T, Ewalt KL;
XX
XX WPI; 2001-281974/29.
XX
XX
XX Recombinant AB5B subunit protein comprising a mutation that alters the
PT number residues available for chemical modification, useful for
PT covalently linking to an immunogen or vaccine which can be used for
PT treating autoimmune diseases -
XX
XX
XX Disclosure; Page 70; 78pp; English.
XX
XX The invention relates to a recombinant AB5B subunit protein comprising at
CC least one mutation, where the mutation alters the number of amino acid
CC residues available for chemical modification as compared to a wild type
CC AB5B subunit protein, and where the recombinant protein retains an
CC effective target ligand binding affinity. A recombinant AB5B subunit
CC protein such as cholera toxin B protein (CTB) can be specifically
CC covalently linked at lysines or cysteines to an immunogen or vaccine.
CC Recombinant mutant CTB (rCTB) can be used in the treatment of autoimmune
CC diseases e.g. rheumatoid arthritis, encephalomyelitis (or other neuron
CC demyelinating diseases) and diabetes. The rCTB or other B subunits of the
CC invention can also be used to induce tolerance to infection, e.g.
CC parasitic infection. The present sequence represents a variant of the
CC E. coli heat labile toxin B (LTB) protein.
XX
XX
XX Sequence 124 AA;
SQ

Query Match 100.0%; Score 107; DB 22; Length 124;
Best Local Similarity 100.0%; Pred. No. 2.1e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GETFQVEVPSQHIDSOKKAI 21
Db 66 getfqvevpsqghidsqkai 86
|||||

RESULT 10
AAB62375 standard; Protein; 124 AA.
XX
XX AAB62375;
XX
XX 29-JUN-2001 (first entry)
XX
XX E. coli LTB protein variant (GI: 145833).
XX
XX
XX AB5B subunit protein; mutation; recombinant; cholera toxin B; CTB;
KM immunogen; autoimmune disease; neuron demyelinating disease; rCTB; LTB;
KM infection; parasitic; immunosuppressive; antiarthritic; antineumatic;
KM antidiabetic; neuroprotective; vaccine; heat labile toxin B; variant.
XX
XX Escherichia coli.
XX
XX WO200127144-A2.
XX
XX 19-APR-2001.
XX
XX 05-OCT-2000; 2000WO-US27607.
XX
XX 08-OCT-1999; 99US-0158561.
XX
XX (ACTI-) ACTIVE BIOTECH AB.
XX
XX Handley HH, Haaparanta T, Ewalt KL;
XX
XX

XX WPI, 2001-281974/29.
XX Recombinant AB5B subunit protein comprising a mutation that alters the
XX number residues available for chemical modification, useful for
XX covalently linking to an immunogen or vaccine which can be used for
XX treating autoimmune diseases -
XX
XX Disclosure: Page 71; 78pp; English.
XX The invention relates to a recombinant AB5B subunit protein comprising at
XX least one mutation, where the mutation alters the number of amino acid
XX residues available for chemical modification as compared to a wild type
XX AB5B subunit protein, and where the recombinant protein retains an
XX effective target ligand binding affinity. A recombinant AB5B subunit
XX protein such as cholera toxin B protein (CTB) can be specifically
XX covalently linked at lysines or cysteines to an immunogen or vaccine.
XX Recombinant mutant CTB (rCTB) can be used in the treatment of autoimmune
XX diseases e.g. rheumatoid arthritis, encephalomyelitis (or other neuron
XX demyelinating diseases) and diabetes. The rCTB or other B subunits of the
XX invention can also be used to induce tolerance to infection, e.g.
XX parasitic infection. The present sequence represents a variant of the
XX E. coli heat labile toxin B (LTB) protein.
XX
XX Sequence 124 AA:

Query Match 100.0%; Score 107; DB 22; Length 124;
Best Local Similarity 100.0%; Pred. No. 2.1e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GEFQVEVPGSQHDSQKRAI 21
|||
DB 66 getfgyevpqsghdsqkai 86

RESULT 11

AAB62380
ID AAB62380 standard; Protein; 124 AA.

XX AAB62380;

DT 29-JUN-2001 (first entry)

DE E. coli LTB protein variant (GI: 69630).

XX AB5B subunit protein; mutation; recombinant; cholera toxin B; CTB;
XX immunogen; autoimmune disease; neuron demyelinating disease; rCTB; LTB;
XX infection; parasitic; immunosuppressive; antiarthritic; antirheumatic;
XX antidiabetic; neuroprotective; vaccine; heat labile toxin B; variant.
XX
XX Escherichia coli.
XX
XX WO200127144-A2.

PD 19-APR-2001.

PF 05-OCT-2000; 2000WO-US27607.

PR 08-OCT-1999; 99US-0158561.

PA (ACTI-) ACTIVE BIOTECH AB.

PI Handley HH, Haaparanta T, Ewalt KL;

DR WPI, 2001-281974/29.

XX Recombinant AB5B subunit protein comprising a mutation that alters the
XX number residues available for chemical modification, useful for
XX covalently linking to an immunogen or vaccine which can be used for
XX treating autoimmune diseases -
XX
XX Disclosure: Page 73; 78pp; English.

XX The invention relates to a recombinant AB5B subunit protein comprising at
XX least one mutation, where the mutation alters the number of amino acid
XX residues available for chemical modification as compared to a wild type
XX AB5B subunit protein, and where the recombinant protein retains an
XX effective target ligand binding affinity. A recombinant AB5B subunit
XX protein such as cholera toxin B protein (CTB) can be specifically
XX covalently linked at lysines or cysteines to an immunogen or vaccine.
XX Recombinant mutant CTB (rCTB) can be used in the treatment of autoimmune
XX diseases e.g. rheumatoid arthritis, encephalomyelitis (or other neuron
XX demyelinating diseases) and diabetes. The rCTB or other B subunits of the
XX invention can also be used to induce tolerance to infection, e.g.
XX parasitic infection. The present sequence represents a variant of the
XX E. coli heat labile toxin B (LTB) protein.
XX
XX Sequence 124 AA:

Query Match 100.0%; Score 107; DB 22; Length 124;
Best Local Similarity 100.0%; Pred. No. 2.1e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GEFQVEVPGSQHDSQKRAI 21
|||
DB 66 getfgyevpqsghdsqkai 86

RESULT 12

AAB73241
ID AAB73241 standard; Protein; 134 AA.

XX AAB73241;

DT 14-MAY-2001 (first entry)

DE Recombinant exotoxin protein variant LTBP.

XX Exotoxin mucosal cell binding motif; nucleic acid delivery;
XX nucleic acid affinity domain; heat-labile enterotoxin.
XX
XX Unidentified.

PN WO200111960-A1.

PD 22-FEB-2001.

PF 18-AUG-2000; 2000WO-US22715.

PR 18-AUG-1999; 99US-0149294.

PA (AGRI-) AGRIVAX INC.

PI Welter LM;

DR WPI, 2001-211103/21.

DR N-PSDB; AAF75712.

XX Novel exotoxin protein variant useful as protein carrier for
XX facilitating gene delivery, comprises a mucosal cell binding motif of
XX an exotoxin and a nucleic acid affinity domain -
XX
XX Example 1; Fig 9; 57pp; English.

XX The present invention relates to recombinant exotoxin protein variants,
XX which comprise an exotoxin mucosal cell binding motif and a nucleic acid
XX affinity domain. The present sequence is one such protein variant. In the
XX present invention the heat-labile enterotoxin (LT) of Escherichia coli
XX was used as an exotoxin. LT is a member of the ADP-ribosylating exotoxin
XX (babe) family. The protein variants are useful for selectively delivering
XX nucleic acid to mucosal cells, for inducing an immune response when the
XX nucleic acid encodes an antigen to which the immune response is desired,
XX for selectively delivering a gene to a mucosal cell, and for achieving
XX expression of a protein in a subject, by administering a composition

CC comprising the protein variant.
XX
SQ Sequence 134 AA;

Query Match 100.0%; Score 107; DB 22; Length 134;
Best Local Similarity 100.0%; Pred. No. 2.3e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GETFOVEVPSQHDSQKAI 21
|
66 getfvevpsqhdsgkai 86

RESULT 13

AAB73242 ID AAB73242 standard; Protein: 142 AA.

XX AAB73242;

XX 14-MAY-2001 (first entry)

DE Recombinant exotoxin protein variant LTBpLh.

XX Exotoxin mucosal cell binding motif; nucleic acid delivery;
KM nucleic acid affinity domain; heat-labile enterotoxin.

XX Unidentified.

XX WO200111960-A1.

XX 22-FEB-2001.

XX 18-AUG-2000; 2000WO-US22715.

XX 18-AUG-1999; 99US-0149294.

XX (AGRI-) AGRIVAX INC.

XX Welter LM;

XX WPI: 2001-211103/21.

XX N-PSDB; AAF75713.

PT Novel exotoxin protein variant useful as protein carrier for
PT facilitating gene delivery, comprises a mucosal cell binding motif of
PT an exotoxin and a nucleic acid affinity domain
XX
XX Example 1; Fig 10; 57pp; English.

XX The present invention relates to recombinant exotoxin protein variants,
CC which comprise an exotoxin mucosal cell binding motif and a nucleic acid
CC affinity domain. The present sequence is one such protein variant. In the
CC present invention the heat-labile enterotoxin (LT) of Escherichia coli
CC was used as an exotoxin. LT is a member of the ADP-ribosylating exotoxin
CC (bAR) family. The protein variants are useful for selectively delivering
CC nucleic acid to mucosal cells, for inducing an immune response when the
CC nucleic acid encodes an antigen to which the immune response is desired,
CC for selectively delivering a gene to a mucosal cell, and for achieving
CC expression of a protein in a subject, by administering a composition
CC comprising the protein variant.

SQ Sequence 142 AA;

Query Match 100.0%; Score 107; DB 22; Length 142;
Best Local Similarity 100.0%; Pred. No. 2.5e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GETFOVEVPSQHDSQKAI 21
|
66 getfvevpsqhdsgkai 86

RESULT 14

AAB73243 ID AAB73243 standard; Protein: 155 AA.

XX AAB73243;

XX 14-MAY-2001 (first entry)

DE Recombinant exotoxin protein variant LTB-P.

XX Exotoxin mucosal cell binding motif; nucleic acid delivery;
KM nucleic acid affinity domain; heat-labile enterotoxin.

XX Unidentified.

XX WO200111960-A1.

XX 22-FEB-2001.

XX 18-AUG-2000; 2000WO-US22715.

XX 18-AUG-1999; 99US-0149294.

XX (AGRI-) AGRIVAX INC.

XX Welter LM;

XX WPI: 2001-211103/21.

XX N-PSDB; AAF75714.

PT Novel exotoxin protein variant useful as protein carrier for
PT facilitating gene delivery, comprises a mucosal cell binding motif of
PT an exotoxin and a nucleic acid affinity domain
XX
XX Example 1; Fig 11; 57pp; English.

XX The present invention relates to recombinant exotoxin protein variants,
CC which comprise an exotoxin mucosal cell binding motif and a nucleic acid
CC affinity domain. The present sequence is one such protein variant. In the
CC present invention the heat-labile enterotoxin (LT) of Escherichia coli
CC was used as an exotoxin. LT is a member of the ADP-ribosylating exotoxin
CC (bAR) family. The protein variants are useful for selectively delivering
CC nucleic acid to mucosal cells, for inducing an immune response when the
CC nucleic acid encodes an antigen to which the immune response is desired,
CC for selectively delivering a gene to a mucosal cell, and for achieving
CC expression of a protein in a subject, by administering a composition
CC comprising the protein variant.

SQ Sequence 155 AA;

Query Match 100.0%; Score 107; DB 22; Length 155;
Best Local Similarity 100.0%; Pred. No. 2.8e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GETFOVEVPSQHDSQKAI 21
|
66 getfvevpsqhdsgkai 86

RESULT 15

AAB73244 ID AAB73244 standard; Protein: 163 AA.

XX AAB73244;

XX 14-MAY-2001 (first entry)

DE Recombinant exotoxin protein variant LTB-P.

XX Exotoxin mucosal cell binding motif; nucleic acid delivery;
KM nucleic acid affinity domain; heat-labile enterotoxin; ds.

XX Unidentified.
 OS
 XX
 PN WO20011960-A1.
 XX
 PD 22-FEB-2001.
 XX
 PF 18-AUG-2000; 2000WO-US22715.
 XX
 PR 18-AUG-1999; 99US-0149294.
 XX
 PA (AGRI-) AGRIVAX INC.
 XX
 PI Welter LM;
 XX
 DR WPI; 2001-211103/21.
 XX N-PSDB; AAF5715.
 PT Novel exotoxin protein variant useful as protein carrier for
 PF facilitating gene delivery, comprises a mucosal cell binding motif of
 PT an exotoxin and a nucleic acid affinity domain
 XX
 PS Example 1; Fig 12; 57pp; English.
 CC The present invention relates to recombinant exotoxin protein variants,
 CC which comprise an exotoxin mucosal cell binding motif and a nucleic acid
 CC affinity domain. The present sequence is one such protein variant. In the
 CC present invention the heat-labile enterotoxin (LT) of *Escherichia coli*
 CC was used as an exotoxin. LT is a member of the ADP-ribosylating exotoxin
 CC (bare) family. The protein variants are useful for selectively delivering
 CC nucleic acid to mucosal cells, for inducing an immune response when the
 CC nucleic acid encodes an antigen to which the immune response is desired,
 CC for selectively delivering a gene to a mucosal cell, and for achieving
 CC expression of a protein in a subject, by administering a composition
 CC comprising the protein variant.
 SQ Sequence 163 AA;

Query Match 100.0%; Score 107; DB 22; Length 163;
 Best Local Similarity 100.0%; Pred. No. 3e-10;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GEPFOVEVPGSOHDSQKRAI 21
 ||||||||||||||||
 Db 66 getfgvevpsqhdsgkrai 86

Search completed: July 3, 2002, 08:54:57
 Job time: 287 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 3, 2002, 08:55:39 ; Search time 34.88 Seconds

(Without alignments)
14.706 Million cell updates/sec

Title: US-09-786-648-5

Perfect score: 107

Sequence: 1 GETFQVEVPGSGHDSOKKAI 21

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

231628

ALIGNMENTS

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA: *
1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep.*
2: /cgn2_6/ptodata/2/1aa/5E.COMB.pep.*
3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep.*
4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep.*
5: /cgn2_6/ptodata/2/1aa/PCTUS.COMB.pep.*
6: /cgn2_6/ptodata/2/1aa/backfile1.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	107	100.0	93	2	US-08-292-968-26 Sequence 26, Appl
2	107	100.0	93	2	US-08-467-974-26 Sequence 26, Appl
3	107	100.0	93	2	US-08-467-536-26 Sequence 26, Appl
4	107	100.0	93	3	US-08-467-976-26 Sequence 26, Appl
5	107	100.0	93	4	US-09-082-514-26 Sequence 26, Appl
6	107	100.0	371	2	US-08-829-026A-6 Sequence 6, Appl
7	101	94.4	102	3	US-08-952-337-5 Sequence 5, Appl
8	101	94.4	102	3	US-08-952-337-6 Sequence 6, Appl
9	101	94.4	103	2	US-08-472-171-2 Sequence 2, Appl
10	101	94.4	103	2	US-08-894-526-2 Sequence 2, Appl
11	101	94.4	103	2	US-09-013-047-2 Sequence 2, Appl
12	101	94.4	103	4	US-09-374-597-2 Sequence 2, Appl
13	101	94.4	103	4	US-09-191-852-21 Sequence 2, Appl
14	101	94.4	103	5	PCT-US95-13376-21 Sequence 21, Appl
15	101	94.4	133	3	US-08-952-337-1 Sequence 1, Appl
16	101	94.4	123	3	US-08-952-337-2 Sequence 2, Appl
17	101	94.4	124	2	US-08-747-410-2 Sequence 2, Appl
18	89	83.2	124	1	US-08-449-045C-4 Sequence 4, Appl
19	89	83.2	124	2	US-08-435-605A-12 Sequence 12, Appl
20	89	83.2	124	6	5223610-3 Patent No. 5223610
21	54	50.5	448	4	US-08-878-989-2 Sequence 2, Appl
22	54	50.5	448	4	US-09-272-796-2 Sequence 2, Appl
23	54	50.5	508	4	US-09-344-700-4 Sequence 4, Appl
24	43	40.2	1810	5	PCT-US95-11684-4 Sequence 4, Appl
25	42	39.3	855	2	US-09-027-337-2 Sequence 2, Appl
26	41	38.3	1022	1	US-08-271-364A-8 Sequence 8, Appl
27	41	38.3	1022	2	US-08-222-715B-27 Sequence 27, Appl

28 40.5 37.9 856 3 US-08-709-784-2 Sequence 2, Appl
29 40.5 37.9 862 2 US-08-209-521-23 Sequence 23, Appl
30 40.5 37.9 862 2 US-08-209-521-30 Sequence 30, Appl
31 40.5 37.9 862 4 US-09-059-461-2 Sequence 2, Appl
32 40.5 37.9 862 4 US-08-961-810-133 Sequence 133, App
33 40.5 37.9 862 4 US-08-352-902D-133 Sequence 2, Appl
34 40 37.4 774 3 US-08-902-632-2 Sequence 1, Appl
35 40 37.4 774 3 US-09-073-354-1 Sequence 1, Appl
36 40 37.4 774 3 US-08-656-005A-1 Sequence 1, Appl
37 40 37.4 774 4 US-09-073-259-1 Sequence 1, Appl
38 40 37.4 774 4 US-09-363-095-1 Sequence 1, Appl
39 40 37.4 774 4 US-09-418-027-1 Sequence 1, Appl
40 40 37.4 778 2 US-08-906-925-4 Sequence 4, Appl
41 40 37.4 779 5 US-08-375-134-12 Sequence 12, Appl
42 40 37.4 779 5 PCT-US95-15263-12 Sequence 12, Appl
43 39.5 36.9 1090 4 US-09-346-237-5 Sequence 5, Appl
44 39 36.4 75 1 US-08-350-884-35 Sequence 35, Appl
45 39 36.4 75 1 US-08-709-173-35 Sequence 35, Appl

RESULT 1
US-08-292-968-26
Sequence 26, Application US/08292968
Patent No. 5856122
GENERAL INFORMATION:
APPLICANT: READ, Randy J.
APPLICANT: STEIN, Penelope E.
APPLICANT: COCKLE, Stephen A.
APPLICANT: OOMEN, Raymond P.
APPLICANT: LOOSMORE, Sheena
APPLICANT: KLEIN, Michel H.
APPLICANT: ARMSTRONG, Glen D.
APPLICANT: HAZES, Bart
TITLE OF INVENTION: MODIFICATION OF PERTUSSIS TOXIN
NUMBER OF SEQUENCES: 46
CORRESPONDENCE ADDRESS:
ADDRESSEE: Slim & Mcburney
STREET: Suite 701, 330 University Avenue
CITY: Toronto
STATE: Ontario
COUNTRY: Canada
ZIP: M5G 1R7
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/292,968
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/110,947
FILING DATE: 24-AUG-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/251,121
FILING DATE: 31-MAY-1994
ATTORNEY/AGENT INFORMATION:
NAME: STEWART, Michael I.
REGISTRATION NUMBER: 24,973
REFERENCE/DOCKET NUMBER: 1038-388
TELECOMMUNICATION INFORMATION:
TELEPHONE: (416) 595-1155
TELEFAX: (416) 595-1163
INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 93 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear

US-08-292-968-26

Query Match : 100.0%; Score 107; DB 2; Length 93;
Best Local Similarity 100.0%; Pred. No. 2.7e-11;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GETFOVEVPGSOHIDSOKKAI 21
|||||

DB 35 GETFOVEVPGSOHIDSOKKAI 55

RESULT 2

US-08-467-974-26
; Sequence 26, Application US/08467974
; Patent No. 5965385

GENERAL INFORMATION:

APPLICANT: READ, Randy J.

APPLICANT: COCKLE, Stephen E.

APPLICANT: OOMEN, Raymond P.

APPLICANT: LOOSMORE, Sheena

APPLICANT: KLEIN, Michael H.

APPLICANT: ARMSTRONG, Glen D.

TITLE OF INVENTION: MODIFICATION OF PERTUSSIS TOXIN

NUMBER OF SEQUENCES: 46

CORRESPONDENCE ADDRESS:

ADDRESSEE: Sim & McBurney

STREET: Suite 701, 330 University Avenue

CITY: Toronto

STATE: Ontario

COUNTRY: Canada

ZIP: M5G 1R7

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: IBM PC compatible

SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/467,974

FILING DATE:

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/467,536

FILING DATE: 06-JUN-1995

APPLICATION NUMBER: US 08/292,968

FILING DATE: 22-AUG-1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/251,121

FILING DATE: 31-MAY-1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/110,947

FILING DATE: 24-AUG-1993

ATTORNEY/AGENT INFORMATION:

NAME: STEWART, Michael I.

REGISTRATION NUMBER: 24,973

REFERENCE/DOCKET NUMBER: 1038-454 MIS:vg

TELECOMMUNICATION INFORMATION:

TELEPHONE: (416) 595-1155

TELEFAX: (416) 595-1163

INFORMATION FOR SEQ ID NO: 26:

SEQUENCE CHARACTERISTICS:

LENGTH: 93 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

US-08-467-974-26

QY 1 GETFOVEVPGSOHIDSOKKAI 21
|||||

DB 35 GETFOVEVPGSOHIDSOKKAI 55

RESULT 3

US-08-467-536-26
; Sequence 26, Application US/08467536
; Patent No. 5977304

GENERAL INFORMATION:

APPLICANT: READ, Randy J.

APPLICANT: COCKLE, Stephen E.

APPLICANT: OOMEN, Raymond P.

APPLICANT: LOOSMORE, Sheena

APPLICANT: KLEIN, Michael H.

APPLICANT: ARMSTRONG, Glen D.

TITLE OF INVENTION: MODIFICATION OF PERTUSSIS TOXIN

NUMBER OF SEQUENCES: 46

CORRESPONDENCE ADDRESS:

ADDRESSEE: Sim & McBurney

STREET: Suite 701, 330 University Avenue

CITY: Toronto

STATE: Ontario

COUNTRY: Canada

ZIP: M5G 1R7

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: IBM PC compatible

SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/467,536

FILING DATE: 06-JUN-1995

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/292,968

FILING DATE: 22-AUG-1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/251,121

FILING DATE: 31-MAY-1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/110,947

FILING DATE: 24-AUG-1993

ATTORNEY/AGENT INFORMATION:

NAME: STEWART, Michael I.

REGISTRATION NUMBER: 24,973

REFERENCE/DOCKET NUMBER: 1038-455 MIS:vg

TELECOMMUNICATION INFORMATION:

TELEPHONE: (416) 595-1155

TELEFAX: (416) 595-1163

INFORMATION FOR SEQ ID NO: 26:

SEQUENCE CHARACTERISTICS:

LENGTH: 93 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

US-08-467-536-26

Query Match : 100.0%; Score 107; DB 2; Length 93;
Best Local Similarity 100.0%; Pred. No. 2.7e-11;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GETFOVEVPGSOHIDSOKKAI 21
|||||

DB 35 GETFOVEVPGSOHIDSOKKAI 55

RESULT 4

US-08-467-976-26

; Sequence 26, Application US/08467976
; Patent No. 6018022
; GENERAL INFORMATION:
; APPLICANT: READ, Randy J.
; APPLICANT: STEIN, Penelope E.
; APPLICANT: COCKLE, Stephen A.
; APPLICANT: OOMEN, Raymond P.
; APPLICANT: LOOSMORE, Sheena
; APPLICANT: KLEIN, Michel H.
; APPLICANT: ARMSTRONG, Glen D.
; TITLE OF INVENTION: MODIFICATION OF PERTUSSIS TOXIN
; NUMBER OF SEQUENCES: 46
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: Suite 701, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/467,976
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/292,968
; FILING DATE: 22-AUG-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/251,121
; FILING DATE: 31-MAY-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/110,947
; FILING DATE: 24-AUG-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: STEWART, Michael I.
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-453 MIS:vg
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; INFORMATION FOR SEQ ID NO: 26:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 93 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-467-976-26

Query Match 100.0%; Score 107; DB 3; Length 93;
Best Local Similarity 100.0%; Pred. No. 2.7e-11;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GETFOVEVPGSQHIDSOKKAI 21
Db 35 GETFOVEVPGSQHIDSOKKAI 55

RESULT 5
US-09-082-514-26
; Sequence 26, Application US/09082514
; Patent No. 6168928
; GENERAL INFORMATION:
; APPLICANT: READ, Randy J.
; APPLICANT: STEIN, Penelope E.
; APPLICANT: COCKLE, Stephen A.
; APPLICANT: OOMEN, Raymond P.
; APPLICANT: KLEIN, Michel H.

; APPLICANT: ARMSTRONG, Glen D.
; APPLICANT: HAZES, Bart
; TITLE OF INVENTION: MODIFICATION OF PERTUSSIS TOXIN
; NUMBER OF SEQUENCES: 46
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: 6th Floor, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/082,514
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/292,968
; FILING DATE: 24-AUG-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: STEWART, Michael I.
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-810
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; INFORMATION FOR SEQ ID NO: 26:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 93 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-09-082-514-26

Query Match 100.0%; Score 107; DB 4; Length 93;
Best Local Similarity 100.0%; Pred. No. 2.7e-11;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GETFOVEVPGSQHIDSOKKAI 21
Db 35 GETFOVEVPGSQHIDSOKKAI 55

RESULT 6
US-08-829-026A-6
; Sequence 6, Application US/08829026A
; Patent No. 5837825
; GENERAL INFORMATION:
; APPLICANT: Melnersmann, Richard J.
; APPLICANT: Khoury, Christian A.
; TITLE OF INVENTION: Campylobacter Jejuni Flagellin-Escherichia Coli LT-B Fusion
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Janelle S. Graeter
; STREET: Room 411, Building 005, BARC-W
; CITY: Beltsville
; STATE: MD
; COUNTRY: USA
; ZIP: 20705
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/829,026A
; FILING DATE: 18-AUG-1997
; CLASSIFICATION: 435

```

?      ATTORNEY/AGENT INFORMATION:
?      NAME:  Greeter, Janelle S.
?      REGISTRATION NUMBER:  35,024
?      REFERENCE/DOCKET NUMBER:  0106,97
?      TELECOMMUNICATION INFORMATION:
?      TELEPHONE:  301-504-5676
?      TELEFAX:  301-504-5060
?      INFORMATION FOR SEO ID NO:  6:
?      SEQUENCE CHARACTERISTICS:
?      LENGTH:  371 amino acids
?      type:  amino acid
?      TOPOLOGY:  linear
?      MOLECULE TYPE:  protein
?      US-08-829-026A-6

```

Query Match	100.0%;	Score 107;	DB 2;	Length 371;
Best Local Similarity	100.0%;	Pred. No. 1.5e-10;		
Matches 21; Conservative	0;	Mismatches	0;	Indels 0; Gaps 0

```
QY 1 GETFOVEVPGSQHIDSQKKA 21
    |||||
Db 49 GETFOVEVPGSQHIDSQKKA 69
```

```

RESULT 7
US-08-952-337-5
: Sequence 5, Application US/08952337
: Patent No. 6019973
: GENERAL INFORMATION:
: APPLICANT: Holmgren, Jan
: APPLICANT: Lebens, Michael R.
: TITLE OF INVENTION: HYBRID MOLECULES BETWEEN HEAT-LABILE
: TITLE OF INVENTION: ENTEROTOXIN AND CHOLERA TOXIN B SUBUNITS
: FILE REFERENCE: 3846/00758
: CURRENT APPLICATION NUMBER: US/08/952,337
: CURRENT FILING DATE: 1998-01-05
: EARLIER APPLICATION NUMBER: PCT/SE96/00570
: EARLIER FILING DATE: 1996-05-02
: EARLIER APPLICATION NUMBER: SE 9501682-0
: EARLIER FILING DATE: 1995-05-05
: NUMBER OF SEQ ID NOS: 6
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 5
: LENGTH: 102
: TYPE: PRT
: ORGANISM: Vibrio cholerae
US-08-952-337-5

```

Query Match	94.48;	Score 101;	DB 3;	Length 102;
Best Local Similarity	95.28;	Pred. No. 3.3e-10;		
Matches 20; Conservative	0;	Mismatches 1;	Indels 0;	Gaps 0;

QY 1 GETTQVEVPGSQHIDSQKKAI 21
| | | | | | | | | | |
Db 44 GATFQVEVPKGSQHIDSQKKAI 64

RESULT 8
 US-08-952-337-6
 : Sequence 6, Application US/08952337
 : Patent No. 6019973
 :
 : GENERAL INFORMATION:
 :
 : APPLICANT: Holmgren, Jan
 : APPLICANT: Lebens, Michael R.
 :
 : TITLE OF INVENTION: HYBRID MOLECULES BETWEEN HEAT-LABILE
 : TITLE OF INVENTION: ENTEROTOXIN AND CHOLERA TOXIN B SUBUNITS
 : FILE REFERENCE: 3846/00758
 :
 : CURRENT APPLICATION NUMBER: US/08/952.337
 :
 : CURRENT FILING DATE: 1998-01-05
 :
 : EARLIER APPLICATION NUMBER: PCT/SE96/00570
 :
 : EARLIER FILING DATE: 1996-05-02

```

1 EARLIER APPLICATION NUMBER: SE 9501682-0
2 EARLIER FILING DATE: 1995-05-05
3 NUMBER OF SEQ. ID NOS: 6
4 SOFTWARE: FASTSEQ for Windows Version 3.
5 SEQ. ID NO. 6
6 LENGTH: 102
7 TYPE: PRN
8 ORGANISM: Escherichia coli
9
10 US-08-952-337-6

```

query Match	94.48;	Score 101;	DB 3;	Length 102;
Best Local Similarity	95.28;	Pred. No. 3.3e-10;		
Matches 20;	Conservative 0;	Mismatches 1;	Indels 0;	Gaps 0;

```
QY 1 GETFQVEVPGSQHIDSQKAI 21
    | | | | | | | | | | | | |
Db 44 GATFQVEVPGSQHIDSQKAI 64
```

```

RESULT          9
US-08-472-171-2
? Sequence 2, Application US/08472171
? Patent No. 5932714
? GENERAL INFORMATION:
? APPLICANT:  Loosmore, Sheena M.
? APPLICANT:  Yacoub, Reza K.
? APPLICANT:  Zealey, Gavin R.
? APPLICANT:  Klehn, Michael H.
? TITLE OF INVENTION:  Expression Of Gene Products From
? TITLE OF INVENTION:  Genetically Manipulated Strains Of Bordetella
? NUMBER OF SEQUENCES:  56
? CORRESPONDENCE ADDRESS:
? ADDRESSEE:  Slim & McBurney
? STREET:  330 University Avenue, Suite 701
? CITY:  Toronto
? STATE:  Ontario
? COUNTRY:  Canada
? ZIP:  M5G 1R7
? COMPUTER READABLE FORM:
? MEDIUM TYPE:  Floppy disk
? COMPUTER:  IBM PC compatible
? OPERATING SYSTEM:  PC-DOS/MS-DOS
? SOFTWARE:  PatentIn Release #1.0, Version #1.30
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER:  US/08/472,171
? FILING DATE:  07-JUN-1995
? CLASSIFICATION:  435
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER:  US 08/393,334
? FILING DATE:  23-FEB-1995
? ATTORNEY/AGENT INFORMATION:
? NAME:  Stewart, Michael I.
? REGISTRATION NUMBER:  24,973
? REFERENCE/DOCKET NUMBER:  1038-507 MIS:vg
? TELECOMMUNICATION INFORMATION:
? TELEPHONE:  416-595-1155
? TELEFAX:  416-595-1163
? TELEX:  065-24567 Simbas
? INFORMATION FOR SEQ ID NO:  2:
? SEQUENCE CHARACTERISTICS:
? LENGTH:  103 amino acids
? TYPE:  amino acid
? TOPOLOGY:  linear
? MOLECULE TYPE:  protein
? US-08-472-171-2

```

Query Match	94.4%	Score 101;	DB 2;	Length 103;
Best Local Similarity	95.2%	Pred. No. 3.3e-10;		
Matches 20; Conservative	0;	Mismatches 1;	Indels 0;	Gaps 0;

QY 1 GETFOVEVPGSQHI DSQKAI 21

Db 45 GATFOVEYPSQSHIDSOKKAI 65

RESULT 10

US-08-894-526-2
Sequence 2, Application US/08894526
Patent No. 5942418

GENERAL INFORMATION:

APPLICANT: Loosmore, Sheena M

APPLICANT: Yacoob, Reza K

APPLICANT: Zealey, Gavin R

APPLICANT: Klein, Michel H

TITLE OF INVENTION: EXPRESSION OF GENE PRODUCTS FROM GENETICALLY MANIPULATED STRAINS OF BORDETTELLA

NUMBER OF SEQUENCES: 57

CORRESPONDENCE ADDRESS:

ADDRESSEE: Sim & McBurney

STREET: 6th Floor, 330 University Avenue

CITY: Toronto

STATE: Ontario

COUNTRY: Canada

ZIP: M5G 1R7

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/894,526

FILING DATE: 01-DEC-1997

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Stewart, Michael I

REGISTRATION NUMBER: 24,973

REFERENCE/DOCKET NUMBER: 1038-724 MIS:jb

TELECOMMUNICATION INFORMATION:

TELEPHONE: (416) 595-1155

TELEFAX: (416) 595-1163

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 103 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-894-526-2

Query Match 94.4% Score 101; DB 2; Length 103;
Best Local Similarity 95.2% Pred. No. 3.3e-10;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GETFOVEYPSQSHIDSOKKAI 21

Db 45 GATFOVEYPSQSHIDSOKKAI 65

RESULT 11

US-09-013-047-2
Sequence 2, Application US/09013047
Patent No. 5998168

GENERAL INFORMATION:

APPLICANT: Loosmore, Sheena M.

APPLICANT: Yacoob, Reza K.

APPLICANT: Zealey, Gavin R.

APPLICANT: Klein, Michel H.

TITLE OF INVENTION: Expression of Gene Products From Genetically Manipulated Strains of Bordetella

NUMBER OF SEQUENCES: 56

CORRESPONDENCE ADDRESS:

ADDRESSEE: Sim & McBurney

STREET: 330 University Avenue, 6th Floor

CITY: Toronto

STATE: Ontario
COUNTRY: Canada

ZIP: M5G 1R7

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/013,047

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/472,171

FILING DATE: 07-JUN-1995

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/393,334

FILING DATE: 23-FEB-1995

ATTORNEY/AGENT INFORMATION:

NAME: Stewart, Michael I.

REGISTRATION NUMBER: 24,973

REFERENCE/DOCKET NUMBER: 1038-507 MIS:y9

TELECOMMUNICATION INFORMATION:

TELEPHONE: 416-595-1155

TELEFAX: 416-595-1163

TELEX: 065-24567 Simbas

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 103 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-09-013-047-2

Query Match 94.4% Score 101; DB 2; Length 103;
Best Local Similarity 95.2% Pred. No. 3.3e-10;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GETFOVEYPSQSHIDSOKKAI 21

Db 45 GATFOVEYPSQSHIDSOKKAI 65

RESULT 12

US-09-374-597-2
Sequence 2, Application US/09374597
Patent No. 6140082

GENERAL INFORMATION:

APPLICANT: Loosmore, Sheena M.

APPLICANT: Yacoob, Reza K.

APPLICANT: Zealey, Gavin R.

APPLICANT: Klein, Michel H.

TITLE OF INVENTION: Expression of Gene Products From Genetically Manipulated Strains of Bordetella

NUMBER OF SEQUENCES: 56

CORRESPONDENCE ADDRESS:

ADDRESSEE: Sim & McBurney

STREET: 330 University Avenue, 6th Floor

CITY: Toronto

STATE: Ontario

COUNTRY: Canada

ZIP: M5G 1R7

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/374,597

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/393,334
FILING DATE: FEBRUARY 23, 1995
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Stewart, Michael I
REGISTRATION NUMBER: 24973
REFERENCE/DOCKET NUMBER: 1038-964
TELECOMMUNICATION INFORMATION:
TELEPHONE: 416-595-1153
TELEFAX: 416-595-1155
TELEX: 065-24567 Simbas
INFORMATION FOR SEQ. ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 103 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-374-597-2

Query Match 94.4%; Score 101; DB 4; Length 103;
Best Local Similarity 95.2%; Pred. No. 3.3e-10;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GATFOVEVPGSOHIDSOKKAI 21
DB 45 GATFOVEVPGSOHIDSOKKAI 65

RESULT 13
US-09-191-852-21
Sequence 21, Application US/09191852
GENERAL INFORMATION:
APPLICANT: Charles J. Arntzen, Hugh S. Mason, and Tariq A. Haq
TITLE OF INVENTION: ORAL IMMUNIZATION WITH TRANSGENIC PLANTS
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fulbright & Jaworski L.L.P.
STREET: 1301 McKinney, Suite 5100
CITY: Houston
STATE: TX
COUNTRY: USA
ZIP: 77010
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/191,852
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/13376
FILING DATE: 24-OCT-1995
APPLICATION NUMBER: 08/817,906
FILING DATE: 04-AUG-1997
ATTORNEY/AGENT INFORMATION:
NAME: Fox, David L.
REGISTRATION NUMBER: 40,612
REFERENCE/DOCKET NUMBER: P015900S1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 713-651-5151
TELEFAX: 713-651-5246
INFORMATION FOR SEQ. ID NO: 21:
SEQUENCE CHARACTERISTICS:
LENGTH: 103 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-09-191-852-21

Query Match 94.4%; Score 101; DB 4; Length 103;
Best Local Similarity 95.2%; Pred. No. 3.3e-10;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GATFOVEVPGSOHIDSOKKAI 21
DB 45 GATFOVEVPGSOHIDSOKKAI 65

RESULT 14
PCT-US95-13376-21
Sequence 21, Application PC/TUS9513376
GENERAL INFORMATION:
APPLICANT: The Texas A&M University System
APPLICANT: 310 Wisenbaker
APPLICANT: College Station, Texas 77843-3369
TITLE OF INVENTION: ORAL IMMUNIZATION WITH TRANSGENIC PLANTS
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pravel, Hewitt, Kimball & Krieger
STREET: 1177 West Loop South, 10th Floor
CITY: Houston
STATE: TX
COUNTRY: USA
ZIP: 77027-9095
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/13376
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/328,716
FILING DATE: 24-OCT-1994
ATTORNEY/AGENT INFORMATION:
NAME: Jones, John W.
REGISTRATION NUMBER: 31,380
REFERENCE/DOCKET NUMBER: 36170/3P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 713-850-0165
TELEFAX: 713-850-0909
INFORMATION FOR SEQ. ID NO: 21:
SEQUENCE CHARACTERISTICS:
LENGTH: 103 amino acids
TYPE: amino acid
TOPOLOGY: linear
PCT-US95-13376-21

Query Match 94.4%; Score 101; DB 5; Length 103;
Best Local Similarity 95.2%; Pred. No. 3.3e-10;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GATFOVEVPGSOHIDSOKKAI 21
DB 45 GATFOVEVPGSOHIDSOKKAI 65

RESULT 15
US-08-952-337-1
Sequence 1, Application US/08952337
Patent No. 6019973
GENERAL INFORMATION:
APPLICANT: Holmgren, Jan
APPLICANT: Lebens, Michael R.
TITLE OF INVENTION: HYBRID MOLECULES BETWEEN HEAT-LABILE
FIDE REFERENCE: 3846/OD758
CURRENT APPLICATION NUMBER: US/08/952,337
CURRENT FILING DATE: 1998-01-05

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 3, 2002, 09:03:21 ; Search time 401.04 Seconds
(without alignments)
18.431 Million cell updates/sec

Title: US-09-786-648-5

Perfect score: 107

Sequence: 1 GERTFYVEFGSHIDSOKKAI 21

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 3502263 seqs, 351980561 residues

Total number of hits satisfying chosen parameters: 3502263

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending Patents AA Main:*

- 1: /cgn2_6/ptodata/2/paa/PCRYUS_COMB.pep.*
- 2: /cgn2_6/ptodata/2/paa/US06_COMB.pep.*
- 3: /cgn2_6/ptodata/2/paa/US07_COMB.pep.*
- 4: /cgn2_6/ptodata/2/paa/US08_COMB.pep.*
- 5: /cgn2_6/ptodata/2/paa/US081_COMB.pep.*
- 6: /cgn2_6/ptodata/2/paa/US082_COMB.pep.*
- 7: /cgn2_6/ptodata/2/paa/US083_COMB.pep.*
- 8: /cgn2_6/ptodata/2/paa/US084_COMB.pep.*
- 9: /cgn2_6/ptodata/2/paa/US085_COMB.pep.*
- 10: /cgn2_6/ptodata/2/paa/US086_COMB.pep.*
- 11: /cgn2_6/ptodata/2/paa/US087_COMB.pep.*
- 12: /cgn2_6/ptodata/2/paa/US088_COMB.pep.*
- 13: /cgn2_6/ptodata/2/paa/US089_COMB.pep.*
- 14: /cgn2_6/ptodata/2/paa/US090_COMB.pep.*
- 15: /cgn2_6/ptodata/2/paa/US091_COMB.pep.*
- 16: /cgn2_6/ptodata/2/paa/US092_COMB.pep.*
- 17: /cgn2_6/ptodata/2/paa/US093_COMB.pep.*
- 18: /cgn2_6/ptodata/2/paa/US094_COMB.pep.*
- 19: /cgn2_6/ptodata/2/paa/US095_COMB.pep.*
- 20: /cgn2_6/ptodata/2/paa/US096_COMB.pep.*
- 21: /cgn2_6/ptodata/2/paa/US097_COMB.pep.*
- 22: /cgn2_6/ptodata/2/paa/US098_COMB.pep.*
- 23: /cgn2_6/ptodata/2/paa/US099_COMB.pep.*
- 24: /cgn2_6/ptodata/2/paa/US100_COMB.pep.*
- 25: /cgn2_6/ptodata/2/paa/US101_COMB.pep.*
- 26: /cgn2_6/ptodata/2/paa/US60_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	107	100.0	21	21	US-09-786-648-5
2	107	100.0	93	5	US-08-110-947-10
3	107	100.0	93	5	US-08-110-947A-26
4	107	100.0	93	6	US-08-251-121-26
5	107	100.0	371	5	US-08-150-305A-3
6	107	100.0	371	11	US-08-784-218-6
7	107	100.0	371	12	US-08-829-026-5

8	101	94.4	21	21	US-09-786-648-4	Sequence 4, Appli
9	101	94.4	103	7	US-08-393-334-2	Sequence 2, Appli
10	101	94.4	103	11	US-08-782-832-15	Sequence 15, Appli
11	101	94.4	103	12	US-08-817-906-21	Sequence 21, Appli
12	101	94.4	103	22	US-09-836-433-14	Sequence 14, Appli
13	101	94.4	116	22	US-09-836-433-20	Sequence 20, Appli
14	101	94.4	119	22	US-09-836-433-22	Sequence 22, Appli
15	101	94.4	124	18	PCT-US99-30747-55	Sequence 55, Appli
16	101	94.4	124	18	US-09-470-124-55	Sequence 55, Appli
17	101	94.4	138	13	US-08-914-479A-2	Sequence 2, Appli
18	101	94.4	138	13	US-08-914-479A-2	Sequence 2, Appli
19	101	94.4	313	21	US-09-756-983-15	Sequence 15, Appli
20	101	94.4	351	21	US-09-756-983-18	Sequence 18, Appli
21	101	94.4	364	21	US-09-736-963-22	Sequence 22, Appli
22	95	88.8	124	1	PCT-US99-30747-57	Sequence 57, Appli
23	95	88.8	124	18	US-09-470-124-57	Sequence 57, Appli
24	89	83.2	124	21	US-09-760-234-7	Sequence 7, Appli
25	89	83.2	382	1	PCT-US01-08582-3	Sequence 3, Appli
26	89	83.2	382	1	PCT-US01-08582-4	Sequence 4, Appli
27	89	83.2	461	14	US-09-051-315-2	Sequence 2, Appli
28	89	83.2	461	18	US-09-423-493-2	Sequence 2, Appli
29	89	83.2	750	18	US-09-402-100-2	Sequence 2, Appli
30	89	83.2	1338	18	US-09-402-100-4	Sequence 4, Appli
31	76	71.0	15	11	US-08-732-371-1	Sequence 1, Appli
32	76	71.0	15	11	US-08-732-371A-1	Sequence 1, Appli
33	62	57.9	12	21	US-09-786-648-3	Sequence 3, Appli
34	54	50.5	286	21	US-09-758-445-415	Sequence 415, App
35	54	50.5	448	21	US-09-769-970-2	Sequence 2, Appli
36	54	50.5	508	19	US-09-563-997-4	Sequence 4, Appli
37	49	45.8	248	18	US-09-451-330-388	Sequence 388, App
38	49	45.8	267	17	US-09-391-631-3104	Sequence 3104, Ap
39	49	45.8	267	20	US-09-689-980-802	Sequence 802, App
40	49	45.8	270	17	US-09-391-631-3103	Sequence 3103, Ap
41	49	45.8	270	20	US-09-689-980-801	Sequence 801, App
42	48	44.9	340	16	US-09-252-691-6923	Sequence 6923, Ap
43	48	44.9	340	16	US-09-252-691C-6923	Sequence 6923, Ap
44	48	44.9	484	16	US-09-252-991A-29252	Sequence 29252, A
45	47	43.9	369	21	US-09-708-427-58248	Sequence 58248, A

ALIGNMENTS

US-09-786-648-5

Sequence 5, Application US/09786648

GENERAL INFORMATION:

APPLICANT: Williams, Neil Andrew

TITLE OF INVENTION: Peptide Fragments of Cholera Toxin B or Enterotoxin B as Vacci

TITLE OF INVENTION: Adjuncts

FILE REFERENCE: 7438

CURRENT APPLICATION NUMBER: US/09/786,648

PRIOR FILING DATE: 2001-03-07

PRIOR APPLICATION NUMBER: PCT/GB99/02970

NUMBER OF SEQ ID NOS: 6

SOFTWARE: MS DOS

SEQ ID NO 5

LENGTH: 21

TYPE: PRT

ORGANISM: E. COLI

FEATURE:

LOCATION: 45...65

OTHER INFORMATION: Isolated or synthetic Etbx beta4-alpha2 loop fragment derivabl

OTHER INFORMATION: porcine E. coli

US-09-786-648-5

Query Match 100.0%, Score 107; DB 21; Length 21;
Best Local Similarity 100.0%; Pred. No. 1.4e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

NAME: STEWART, Michael I.
REGISTRATION NUMBER: 24,973
REFERENCE/DOCKET NUMBER: 1038-335
TELECOMMUNICATION INFORMATION:
TELEPHONE: (416) 595-1155
TELEFAX: (416) 595-1163
INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 93 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-251-121-26

Query Match 100.0%; Score 107; DB 6; Length 93;
Best Local Similarity 100.0%; Pred. No. 9.3e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GETFOVEVPGSQHIDSOKKAI 21
|||||
Db 35 GETFOVEVPGSQHIDSOKKAI 55

RESULT 5
US-08-150-305A-3
Sequence 3, Application US/08150305A
GENERAL INFORMATION:
APPLICANT: Meinersmann, Richard J.
TITLE OF INVENTION: CAMPYLOBACTER JEJUNI
TITLE OF INVENTION: FLAGELLIN-ESCHERICHIA COLI ST-B FUSION PROTEIN
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: Janelle S. Graeter
STREET: Room 413, Building 005, BARC-W
CITY: Beltsville
STATE: MD
COUNTRY: USA
ZIP: 20705
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/150,305A
FILING DATE: 12-NOV-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Graeter, Janelle S.
REGISTRATION NUMBER: 35,024
REFERENCE/DOCKET NUMBER: 0431.92
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-504-5676
TELEFAX: 202-504-5060
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 371 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-150-305A-3

Query Match 100.0%; Score 107; DB 5; Length 371;
Best Local Similarity 100.0%; Pred. No. 5.5e-09;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GETFOVEVPGSQHIDSOKKAI 21
|||||
Db 49 GETFOVEVPGSQHIDSOKKAI 69

RESULT 6
US-08-784-218-6
Sequence 6, Application US/08784218
GENERAL INFORMATION:
APPLICANT: Meinersmann, Richard J.
TITLE OF INVENTION: CAMPYLOBACTER JEJUNI FLAGELLIN-ESCHERICHIA COLI LT-B FUSION
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Janelle S. Graeter
STREET: Room 411, Building 005, BARC-W
CITY: Beltsville
STATE: MD
COUNTRY: USA
ZIP: 20705
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/784,218
FILING DATE: 16-JAN-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Graeter, Janelle S.
REGISTRATION NUMBER: 35,024
REFERENCE/DOCKET NUMBER: 0043.97
TELECOMMUNICATION INFORMATION:
TELEPHONE: 301-504-5676
TELEFAX: 301-504-5060
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 371 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-784-218-6

Query Match 100.0%; Score 107; DB 11; Length 371;
Best Local Similarity 100.0%; Pred. No. 5.5e-09;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GETFOVEVPGSQHIDSOKKAI 21
|||||
Db 49 GETFOVEVPGSQHIDSOKKAI 69

RESULT 7
US-08-829-026-5
Sequence 5, Application US/08829026
GENERAL INFORMATION:
APPLICANT: Meinersmann, Richard J.
TITLE OF INVENTION: CAMPYLOBACTER JEJUNI
TITLE OF INVENTION: FLAGELLIN-ESCHERICHIA COLI LT-B FUSION PROTEIN
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Janelle S. Graeter
STREET: Room 411, Building 005, BARC-W
CITY: Beltsville
STATE: MD
COUNTRY: USA
ZIP: 20705
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/829,026

;; FILING DATE: 18-AUG-1997
;; CLASSIFICATION: 435
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Graeter, Janelle S.
;; REGISTRATION NUMBER: 35,024
;; REFERENCE/DOCKET NUMBER: 0106.97
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 301-504-5676
;; TELEFAX: 301-504-5060
;; INFORMATION FOR SEQ ID NO: 5:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 371 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
US-08-829-026-5

Query Match 100.0%; Score 107; DB 12; Length 371;
Best Local Similarity 100.0%; Pred. No. 5.5e-09;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GETFQVEVPGSQHIDSQKAI 21
Db 49 GETFQVEVPGSQHIDSQKAI 69

RESULT 8
US-09-786-648-4
; Sequence 4, Application US/09786648
; GENERAL INFORMATION:
; APPLICANT: Williams, Neil Andrew
; APPLICANT: Hirst, Timothy Raymond
; TITLE OF INVENTION: Peptide Fragments of Cholera Toxin B or Enterotoxin B as Vaccine
; FILE REFERENCE: 7438
; CURRENT APPLICATION NUMBER: US/09/786,648
; PRIOR FILING DATE: 2001-03-07
; PRIOR APPLICATION NUMBER: PCT/GB99/02970
; PRIOR FILING DATE: 1999-09-07
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: MS DOS
; SEQ ID NO: 4
; LENGTH: 21
; TYPE: PRT
; ORGANISM: E. coli
; FEATURE:
; LOCATION: 45..65
; OTHER INFORMATION: Isolated or synthetic EtxB beta4-alpha2 loop fragment derivable
US-09-786-648-4

Query Match 94.4%; Score 101; DB 21; Length 21;
Best Local Similarity 95.2%; Pred. No. 1.4e-09;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GETFQVEVPGSQHIDSQKAI 21
Db 1 GATFQVEVPGSQHIDSQKAI 21

RESULT 9
US-08-393-334-2
; Sequence 2, Application US/08393334
; GENERAL INFORMATION:
; APPLICANT: Loomore, Sheena M.
; APPLICANT: Yacoub, Reza K.
; APPLICANT: Zealey, Gavin R.
; APPLICANT: Klein, Michel H.
; TITLE OF INVENTION: Expression of Gene Products From
; TITLE OF INVENTION: Genetically Manipulated Strains of Bordetella
; NUMBER OF SEQUENCES: 36

;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Sim & McBurney
;; STREET: 330 University Avenue, Suite 701
;; CITY: Toronto
;; STATE: Ontario
;; COUNTRY: Canada
;; ZIP: M5G 1R7
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/393,334
;; FILING DATE:
;; CLASSIFICATION: 435
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Stewart, Michael I
;; REGISTRATION NUMBER: 24973
;; REFERENCE/DOCKET NUMBER: 1038-417
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 416-595-1155
;; TELEFAX: 416-595-1163
;; TELEEX: 065-24567 Simbas
;; INFORMATION FOR SEQ ID NO: 2:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 103 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
US-08-393-334-2

Query Match 94.4%; Score 101; DB 7; Length 103;
Best Local Similarity 95.2%; Pred. No. 1.1e-08;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GETFQVEVPGSQHIDSQKAI 21
Db 45 GATFQVEVPGSQHIDSQKAI 65

RESULT 10
US-08-782-832-15
; Sequence 15, Application US/08782832
; GENERAL INFORMATION:
; APPLICANT: Arntzen, Charles J.
; APPLICANT: Mason, Hugh S.
; APPLICANT: Haq, Tariq A.
; TITLE OF INVENTION: PRODUCTION OF AN ORALLY IMMUNOGENIC
; TITLE OF INVENTION: BACTERIAL PROTEIN IN TRANSGENIC PLANTS
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pravel, Hewitt, Kimball & Krieger
; STREET: 1177 West Loop South, 10th Floor
; CITY: Houston
; STATE: TX
; COUNTRY: USA
; ZIP: 77027-9095
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/782,832
; FILING DATE: 13-JAN-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/328,716
; FILING DATE: 24-OCT-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Kettelberger, Denise M.


```

; SEQ ID NO 22
;
; LENGTH: 119
;
; TYPE: PR1
;
; ORGANISM: synthetic construct
;
US-09-836-433-22

```


GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 3, 2002, 09:04:12 ; Search time 46.52 Seconds

(Without alignments)
45.673 Million cell updates/sec

Title: US-09-786-648-5

Sequence: 1 GETFOVEYPSQHDSOKKAI 21

Scoring table: BIOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 308740 seqs, 101176262 residues

Total number of hits satisfying chosen parameters: 308740

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : Pending Patents, AA, New.*

1: /cgn2_6/prodata/2/paa/PCT_NEW_COMB.pep.*
2: /cgn2_6/prodata/2/paa/US06_NEW_COMB.pep.*
3: /cgn2_6/prodata/2/paa/US07_NEW_COMB.pep.*
4: /cgn2_6/prodata/2/paa/US08_NEW_COMB.pep.*
5: /cgn2_6/prodata/2/paa/US09_NEW_COMB.pep.*
6: /cgn2_6/prodata/2/paa/US10_NEW_COMB.pep.*
7: /cgn2_6/prodata/2/paa/US60_NEW_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	107	100.0	103	6	US-10-110-364-22
2	107	100.0	123	6	US-10-110-364-20
3	107	100.0	124	6	US-10-110-364-16
4	107	100.0	124	6	US-10-110-364-18
5	107	100.0	124	6	US-10-110-364-23
6	101	94.4	103	6	US-10-110-364-10
7	101	94.4	103	6	US-10-110-364-13
8	101	94.4	105	6	US-10-110-364-12
9	101	94.4	123	6	US-10-110-364-17
10	101	94.4	124	6	US-10-110-364-2
11	101	94.4	124	6	US-10-110-364-6
12	101	94.4	124	6	US-10-110-364-11
13	101	94.4	124	6	US-10-110-364-19
14	101	94.4	124	6	US-10-110-364-21
15	101	94.4	125	6	US-10-110-364-15
16	101	94.4	138	6	US-10-141-627-2
17	95	88.8	103	6	US-10-110-364-8
18	92	86.0	103	6	US-10-110-364-7
19	89	83.2	103	6	US-10-110-364-5
20	89	83.2	124	6	US-10-110-364-4
21	89	83.2	124	6	US-10-110-364-9
22	89	83.2	382	5	US-09-809-033A-3
23	89	83.2	382	5	US-09-809-033A-4
24	45	42.1	84	5	US-09-620-393B-3233
25	45	42.1	138	5	US-09-620-393B-3232
26	45	42.1	211	5	US-09-540-209B-9908

27	45	42.1	549	5	US-09-540-209B-6846	Sequence 6846, Ap
28	44	41.1	371	6	US-10-155-881-29649	Sequence 29649, A
29	44	41.1	890	7	US-60-360-039-3507	Sequence 3507, Ap
30	43.5	40.7	641	5	US-09-573-655B-1493	Sequence 1493, Ap
31	43.5	40.7	641	7	US-60-382-898-321	Sequence 321, App
32	43	40.2	747	7	US-60-360-039-7808	Sequence 7808, Ap
33	43	40.2	1810	4	US-08-793-273C-4	Sequence 4, App1
34	42	39.3	359	6	US-10-138-701-36	Sequence 36, App1
35	42	39.3	492	6	US-10-155-881-18930	Sequence 18930, A
36	42	39.3	782	1	PCT-US02-07826-312	Sequence 312, App
37	42	39.3	782	6	US-10-097-340-312	Sequence 312, App
38	42	39.3	855	1	PCT-US02-09039-2	Sequence 2, App1
39	42	39.3	855	5	US-09-654-600A-2	Sequence 2, App1
40	42	39.3	855	6	US-10-104-271-2	Sequence 2, App1
41	42	39.3	855	6	US-10-099-700A-2	Sequence 2, App1
42	42	39.3	1287	5	US-09-573-655B-1370	Sequence 1370, Ap
43	42	39.3	2359	6	US-10-155-881-28520	Sequence 28520, A
44	41.5	38.8	187	6	US-10-155-881-22303	Sequence 22303, A
45	41.5	38.8	209	6	US-10-155-881-33927	Sequence 33927, A

ALIGNMENTS

RESULT 1
US-10-110-364-22
; Sequence 22, Application US/10110364
; GENERAL INFORMATION:
; APPLICANT: Handley, Harold H.
; APPLICANT: Haaparaanta, Tapio
; APPLICANT: Ewailt, Karla L.
; TITLE OF INVENTION: ABS TOXIN B SUBUNIT MUTANTS WITH ALTERED
; FILE REFERENCE: ACTBIO.004A
; CURRENT APPLICATION NUMBER: US/10/110,364
; PRIOR FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: PCT/US00/27607
; PRIOR FILING DATE: 2000-10-05
; PRIOR APPLICATION NUMBER: 60/158,561
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 22
; LENGTH: 103
; TYPE: PRT
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(103)
; OTHER INFORMATION: LTB variant from NCBI gene bank GI: 494265.
US-10-110-364-22

Query Match 100.0%; Score 107; DB 6; Length 103;
Best Local Similarity 100.0%; Pred. No. 6, 3e-11;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 45 GETFOVEYPSQHDSOKKAI 65
|||||

RESULT 2
US-10-110-364-20
; Sequence 20, Application US/10110364
; GENERAL INFORMATION:
; APPLICANT: Handley, Harold H.
; APPLICANT: Haaparaanta, Tapio
; APPLICANT: Ewailt, Karla L.
; TITLE OF INVENTION: ABS TOXIN B SUBUNIT MUTANTS WITH ALTERED
; FILE REFERENCE: ACTBIO.004A
; CURRENT APPLICATION NUMBER: US/10/110,364

```

; CURRENT FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: PCT/US00/27607
; PRIOR FILING DATE: 2000-10-05
; PRIOR APPLICATION NUMBER: 60/158,561
; PRIOR FILING DATE: 1999-10-08
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 20
; LENGTH: 123
; TYPE: PRF
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(123)
; OTHER INFORMATION: LTB variant from NCBI gene bank GI: 223254.
US-10-110-364-20

```

```

Query Match          100.0%; Score 107; DB 6; Length 123;
Best Local Similarity 100.0%; Pred. No. 7.9e-11;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 GETFOVEVPGSOHIDSOKKAI 21
Db 66 GETFOVEVPGSOHIDSOKKAI 85

```

```

RESULT 3
US-10-110-364-16
; Sequence 16, Application US/10110364
; GENERAL INFORMATION:
; APPLICANT: Handley, Harold H.
; APPLICANT: Haaparanta, Tapio
; APPLICANT: Ewalt, Karla L.
; TITLE OF INVENTION: CHEMICAL CONJUGATION CHARACTERISTICS
; FILE REFERENCE: ACTBIO.004A
; CURRENT FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: PCT/US00/27607
; PRIOR FILING DATE: 2000-10-05
; PRIOR APPLICATION NUMBER: 60/158,561
; PRIOR FILING DATE: 1999-10-08
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16
; LENGTH: 124
; TYPE: PRF
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(124)
; OTHER INFORMATION: LTB variant from NCBI gene bank GI: 1169505.
US-10-110-364-16

```

```

Query Match          100.0%; Score 107; DB 6; Length 124;
Best Local Similarity 100.0%; Pred. No. 7.9e-11;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 GETFOVEVPGSOHIDSOKKAI 21
Db 66 GETFOVEVPGSOHIDSOKKAI 86

```

```

RESULT 4
US-10-110-364-18
; Sequence 18, Application US/10110364
; GENERAL INFORMATION:
; APPLICANT: Handley, Harold H.
; APPLICANT: Haaparanta, Tapio
; APPLICANT: Ewalt, Karla L.
; TITLE OF INVENTION: ABS TOXIN B SUBUNIT MUTANTS WITH ALTERED

```

```

; TITLE OF INVENTION: CHEMICAL CONJUGATION CHARACTERISTICS
; FILE REFERENCE: ACTBIO.004A
; CURRENT APPLICATION NUMBER: US/10/110,364
; CURRENT FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: PCT/US00/27607
; PRIOR FILING DATE: 2000-10-05
; PRIOR APPLICATION NUMBER: 60/158,561
; PRIOR FILING DATE: 1999-10-08
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 18
; LENGTH: 124
; TYPE: PRF
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(124)
; OTHER INFORMATION: LTB variant from NCBI gene bank GI: 145833.
US-10-110-364-18

```

```

Query Match          100.0%; Score 107; DB 6; Length 124;
Best Local Similarity 100.0%; Pred. No. 7.9e-11;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 GETFOVEVPGSOHIDSOKKAI 21
Db 66 GETFOVEVPGSOHIDSOKKAI 86

```

```

RESULT 5
US-10-110-364-23
; Sequence 23, Application US/10110364
; GENERAL INFORMATION:
; APPLICANT: Handley, Harold H.
; APPLICANT: Haaparanta, Tapio
; APPLICANT: Ewalt, Karla L.
; TITLE OF INVENTION: ABS TOXIN B SUBUNIT MUTANTS WITH ALTERED
; FILE REFERENCE: ACTBIO.004A
; CURRENT FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: PCT/US00/27607
; PRIOR FILING DATE: 2000-10-05
; PRIOR APPLICATION NUMBER: 60/158,561
; PRIOR FILING DATE: 1999-10-08
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 23
; LENGTH: 124
; TYPE: PRF
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(124)
; OTHER INFORMATION: LTB variant from NCBI gene bank GI: 69630.
US-10-110-364-23

```

```

Query Match          100.0%; Score 107; DB 6; Length 124;
Best Local Similarity 100.0%; Pred. No. 7.9e-11;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 GETFOVEVPGSOHIDSOKKAI 21
Db 66 GETFOVEVPGSOHIDSOKKAI 86

```

```

RESULT 6
US-10-110-364-10
; Sequence 10, Application US/10110364
; GENERAL INFORMATION:
; APPLICANT: Handley, Harold H.

```

```

1  APPLICANT: Haaparanta, Tapio
2  APPLICANT: Ewatt, Karla L.
3  TITLE OF INVENTION: ABS TOXIN B SUBUNIT MUTANTS WITH ALTERED
4  TITLE OF INVENTION: CHEMICAL CONUGATION CHARACTERISTICS
5  FILE REFERENCE: ACTBIO.004A
6  CURRENT APPLICATION NUMBER: US/10/110.364
7  CURRENT FILING DATE: 2002-04-05
8  PRIOR APPLICATION NUMBER: PCT/US00/27607
9  PRIOR FILING DATE: 2000-10-05
10 PRIOR APPLICATION NUMBER: 60/158,561
11 PRIOR FILING DATE: 1999-10-08
12 NUMBER OF SEQ ID NOS: 31
13 SOFTWARE: FastSeq for Windows Version 4.0
14 SEQ ID NO: 10
15 LENGTH: 103
16 TYPE: PRT
17 ORGANISM: Vibrio cholera
18 FEATURE:
19 NAME/KEY: VARIANT
20 LOCATION: (1)...(103)
21 OTHER INFORMATION: CTB variant from NCBI gene bank GI: 14215111
22 US-10-110-364-10

```

Query Match	94.48;	Score 101;	DB 6;	Length 103;
Best Local Similarity	95.28;	Pred. No. 6.7e-10;		
Matches 20;	Conservative	0;	Mismatches 1;	Indels 0;
			Gaps	0;

```
QY      1 GETFQVEVPGSQHIDSQKAI 21
          | | | | | | | | | | | |
Db      45 GATFQVEVPGSQHIDSQKAI 65
```

```

RESULT 7
US-10-110-364-13
; Sequence 13, Application US/10110364
; GENERAL INFORMATION:
; APPLICANT: Handley, Harold H.
; APPLICANT: Haaparenta, Tapio
; APPLICANT: Ewalt, Karla L.
; TITLE OF INVENTION: ABS TOXIN B SUBUNIT MUTANTS WITH ALTERED
; TITLE OF INVENTION: CHEMICAL CONJUGATION CHARACTERISTICS
; FILE REFERENCE: ACTBIO.004A
; CURRENT APPLICATION NUMBER: US/10-110, 364
; CURRENT FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: PCT/US00/27607
; PRIOR FILING DATE: 2000-10-05
; PRIOR APPLICATION NUMBER: 60/158, 561
; PRIOR FILING DATE: 1999-10-08
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 103
; TYPE: PRT
; ORGANISM: Vibrio cholera
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(103)
; OTHER INFORMATION: CTB variant from NCBI gene bank GI: 14215235
; OTHER INFORMATION: (ogawa 41 R35D).
; US-10-110-364-13

```

Query Match	94.4%	Score 101;	DB 6;	Length 103;
Best Local Similarity	95.2%;	Pred. No. 6.7e-10;		
Matches	20;	Conservative	0;	Mismatches 1;
			Indels	0;
			Gaps	0;

```
QY      1 GETFQVEVPGSQHIDSQKAI 21
          | | | | | | | | | | | |
Db      45 GATFQVEVPGSQHIDSQKAI 65
```

RESULT 8

```

US-10-110-364-12
Sequence 12, Application US/10110364
GENERAL INFORMATION:
APPLICANT: Handley, Harold H.
APPLICANT: Haaparanta, Tapio
APPLICANT: Ewall, Karla L.
TITLE OF INVENTION: ABS TOXIN B SUBUNIT MUTANTS WITH ALTERED
TITLE OF INVENTION: CHEMICAL CONJUGATION CHARACTERISTICS
FILE REFERENCE: ACTBIO.004A
CURRENT APPLICATION NUMBER: US/10/110.364
PRIORITY FILING DATE: 2002-04-05
PRIORITY APPLICATION NUMBER: PCT/US00/27607
PRIORITY FILING DATE: 2000-10-05
PRIORITY APPLICATION NUMBER: 60/158,561
PRIORITY FILING DATE: 1999-10-08
NUMBER OF SEQ ID NOS: 31
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 12
LENGTH: 105
TYPE: PR1
ORGANISM: Vibrio cholera
FEATURE:
NAME/KEY: VARIANT
LOCATION: (1)...(105)
OTHER INFORMATION: CTB variant from NCBI gene bank GI: 2781121 (ogawa)
US-10-110-364-12

```

Query Match	94.48;	Score 101;	DB 6;	Length 105;
Best Local Similarity	95.28;	Pred. No. 6.8e-10;		
Matches 20; Conservative	0;	Mismatches 1;	Indels 0;	Gaps 0;

```
QY 1 GETFOVEVPGSQHIDSQKAI 21
    1 11111111111111111111
Db 46 GATFOVEVPGSQHIDSQKAI 66
```

```

RESULT      9
US-10-110-364-17
; Sequence 17, Application US/10110364
; GENERAL INFORMATION:
; APPLICANT: Handley, Harold H.
; APPLICANT: Haaparanta, Tapio
; APPLICANT: Ewalt, Karla L.
; TITLE OF INVENTION: ABS TOXIN B SUBUNIT MUTANTS WITH ALTERED
; TITLE OF INVENTION: CHEMICAL CONJUGATION CHARACTERISTICS
; FILE REFERENCE: ACTBIO.004A
; CURRENT APPLICATION NUMBER: US/10/110,364
; CURRENT FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: PCT/US00/27607
; PRIOR FILING DATE: 2000-10-05
; PRIOR APPLICATION NUMBER: 60/158,561
; PRIOR FILING DATE: 1999-10-08
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 123
; TYPE: PRT
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(123)
; OTHER INFORMATION: LTB variant from NCBI gene bank GI: 1395122
US-10-110-364-17

```

Query Match	94.48;	Score 101;	DB 6;	Length 123;
Best Local Similarity	95.28;	Pred. No. 8.3e-10;		
Matches 20; Conservative	0;	Mismatches 1;	Indels 0;	Gaps 0;

```

0Y      1 GETFOVEVPGSHIDSQKAI 21
      | | | | | | | | | | | | | |

```

Db 66 GATFQVEVPGSQHIDSOKKAI 86

```
RESULT 10
US-10-110-364-2
; Sequence 2, Application US/10110364
; GENERAL INFORMATION:
; APPLICANT: Handley, Harold H.
; APPLICANT: Haaparanta, Tapio
; APPLICANT: Ewalt, Karla L.
; TITLE OF INVENTION: ABS TOXIN B SUBUNIT MUTANTS WITH ALTERED
; FILE REFERENCE: ACTBIO.004A
; CURRENT FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: PCT/US00/27607
; PRIOR FILING DATE: 2000-10-05
; PRIOR APPLICATION NUMBER: 60/158,561
; PRIOR FILING DATE: 1999-10-08
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 124
; TYPE: PRT
; ORGANISM: Vibrio cholera
US-10-110-364-2
```

Query Match 94.4%; Score 101; DB 6; Length 124;
Best Local Similarity 95.2%; Pred. No. 8.3e-10;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GETFQVEVPGSQHIDSOKKAI 21
Db 66 GATFQVEVPGSQHIDSOKKAI 86

```
RESULT 11
US-10-110-364-6
; Sequence 6, Application US/10110364
; GENERAL INFORMATION:
; APPLICANT: Handley, Harold H.
; APPLICANT: Haaparanta, Tapio
; APPLICANT: Ewalt, Karla L.
; TITLE OF INVENTION: ABS TOXIN B SUBUNIT MUTANTS WITH ALTERED
; FILE REFERENCE: ACTBIO.004A
; CURRENT FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: PCT/US00/27607
; PRIOR FILING DATE: 2000-10-05
; PRIOR APPLICATION NUMBER: 60/158,561
; PRIOR FILING DATE: 1999-10-08
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 124
; TYPE: PRT
; ORGANISM: Vibrio cholera
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(124)
; OTHER INFORMATION: CTB variant from NCBI gene bank GI: 808900.
US-10-110-364-6
```

Query Match 94.4%; Score 101; DB 6; Length 124;
Best Local Similarity 95.2%; Pred. No. 8.3e-10;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GETFQVEVPGSQHIDSOKKAI 21
Db 66 GATFQVEVPGSQHIDSOKKAI 86

RESULT 12
US-10-110-364-11

```
; Sequence 11, Application US/10110364
; GENERAL INFORMATION:
; APPLICANT: Handley, Harold H.
; APPLICANT: Haaparanta, Tapio
; APPLICANT: Ewalt, Karla L.
; TITLE OF INVENTION: ABS TOXIN B SUBUNIT MUTANTS WITH ALTERED
; FILE REFERENCE: ACTBIO.004A
; CURRENT FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: PCT/US00/27607
; PRIOR FILING DATE: 2000-10-05
; PRIOR APPLICATION NUMBER: 60/158,561
; PRIOR FILING DATE: 1999-10-08
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 124
; TYPE: PRT
; ORGANISM: Vibrio cholera
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(124)
; OTHER INFORMATION: CTB variant from NCBI gene bank GI: 48890 (classic)
US-10-110-364-11
```

Query Match 94.4%; Score 101; DB 6; Length 124;
Best Local Similarity 95.2%; Pred. No. 8.3e-10;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GETFQVEVPGSQHIDSOKKAI 21
Db 66 GATFQVEVPGSQHIDSOKKAI 86

```
RESULT 13
US-10-110-364-19
; Sequence 19, Application US/10110364
; GENERAL INFORMATION:
; APPLICANT: Handley, Harold H.
; APPLICANT: Haaparanta, Tapio
; APPLICANT: Ewalt, Karla L.
; TITLE OF INVENTION: ABS TOXIN B SUBUNIT MUTANTS WITH ALTERED
; FILE REFERENCE: ACTBIO.004A
; CURRENT FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: PCT/US00/27607
; PRIOR FILING DATE: 2000-10-05
; PRIOR APPLICATION NUMBER: 60/158,561
; PRIOR FILING DATE: 1999-10-08
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 19
; LENGTH: 124
; TYPE: PRT
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(124)
; OTHER INFORMATION: LTB variant from NCBI gene bank GI: 1648865
US-10-110-364-19
```

Query Match 94.4%; Score 101; DB 6; Length 124;
Best Local Similarity 95.2%; Pred. No. 8.3e-10;

	Matches	20, Conservative	0;	Mismatches	1;	Indels	0;	Gaps	0;
QY	1	GETPVEVPGSQHIDSQKKAI	21						
	1								
Db	66	GATPVEVPGSQHIDSQKKAI	86						

RESULT 14
US-10-110-364-21

```

Sequence 21, Application US/10110364
GENERAL INFORMATION:
APPLICANT: Handley, Harold H.
APPLICANT: Haaparanta, Tapio
APPLICANT: Ewall, Karla L.
TITLE OF INVENTION: ABS TOXIN B SUBUNIT MUTANTS WITH ALTERED
TITLE OF INVENTION: CHEMICAL CONJUGATION CHARACTERISTICS
FILE REFERENCE: AC2BIO.004A
CURRENT APPLICATION NUMBER: US/10/110,364
CURRENT FILING DATE: 2002-04-05
PRIOR APPLICATION NUMBER: PCT/US00/27607
PRIOR FILING DATE: 2000-10-05
PRIOR APPLICATION NUMBER: 60/158,561
PRIOR FILING DATE: 1999-10-08
NUMBER OF SEQ ID NOS: 31
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 21
LENGTH: 124
TYPE: PRT
ORGANISM: Escherichia coli
FEATURE:
NAME/KEY: VARIANT
LOCATION: (1)...(124)
OTHER INFORMATION: LTB variant from NCBI gene bank GI: 408996.
US-10-110-364-21

```

Query Match	94.48%	Score 101	DB 60	Length 124
Best Local Similarity	95.28%	Pred. No. 8.3e-10		
Matches 20	Conservative 0	Mismatches 1	Indels 0	Gaps 0

```
QY 1 GETFQVEVPGSQHIDSQKAI 21
    | | | | | | | | | | | | |
Db 66 GATFQVEVPGSQHIDSQKAI 86
```

RESULT 15
US-10-110-

```

: Sequence 15, Application US/0101364
: GENERAL INFORMATION:
: APPLICANT: Handley, Harold H.
: APPLICANT: Haaparanta, Tapio
: APPLICANT: Ewalt, Karla L.
: TITLE OF INVENTION: Aβ5 TOXIN B SUBUNIT MUTANTS WITH ALTERED
: TITLE OF INVENTION: CHEMICAL CONJUGATION CHARACTERISTICS
: FILE REFERENCE: ACTBIO.000A
: CURRENT APPLICATION NUMBER: US/10/110,364
: CURRENT FILING DATE: 2002-04-05
: PRIOR APPLICATION NUMBER: PCT/US00/27607
: PRIOR FILING DATE: 2000-10-05
: PRIOR APPLICATION NUMBER: 60/158,561
: PRIOR FILING DATE: 1999-10-08
: NUMBER OF SEQ ID NOS: 31
: SOFTWARE: Fastseq for Windows Version 4.0
: SEQ ID NO 15
: LENGTH: 125
: TYPE: PRT
: ORGANISM: Escherichia coli
: FEATURE:
: NAME/KEY: VARIANT
: LOCATION: (1)..(125)
: OTHER INFORMATION: LTB variant from NCBI gene bank GI: 3062900
: US-10-110-364-15

```

query match	94.48;	Score 101;	DB 6;	Length 125;
Best Local Similarity	95.28;	Pred. No. 8.4e-10;		
Matches 20;	Conservative 0;	Mismatches 1;	Indels 0;	Gaps 0;

```
QY 1 GETFQVEVPGSQHIDSQKAI 21
    | | | | | | | | | | | | | |
Db 67 GATFQVEVPGSQHIDSQKAI 87
```

```
Search completed:  July  3, 2002, 09:04:12
Job time: 652 sec
```

[illegible]

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 3, 2002, 08:56:33 ; Search time 46.57 seconds
(without alignments)
43.330 Million cell updates/sec

Title: US-09-786-648-5

Perfect score: 107
Sequence: 1 GETFGVEVPGSGHDSQKKAI 21

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 segs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

PIR-71:*
1: pirl:*
2: pirl:*
3: pirl:*
4: pirl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	107	100.0	124	1	heat-labile enterotoxin B precursor - Escherichia coli
2	95	88.8	124	1	heat-labile enterotoxin B precursor - Escherichia coli
3	49	45.8	255	1	cholera enterotoxin B subunit cistrons
4	47	43.9	392	2	probable peptide c
5	46	43.0	91	2	RAD23 protein homo
6	46	43.0	574	2	hypothetical prote
7	45	42.1	439	2	hypothetical prote
8	45	42.1	2733	2	kinesin heavy chain
9	44	41.1	2633	2	genome polypeptide
10	44	41.1	374	2	conserved hypochet
11	44	41.1	864	2	hypothetical prote
12	44	41.1	2731	1	probable carnitine
13	43.5	40.7	641	2	genome polypeptide
14	43	40.2	128	1	probable receptor
15	43	40.2	154	1	ubiquitin / riboso
16	43	40.2	159	2	ubiquitin / riboso
17	43	40.2	159	2	hypothetical prote
18	43	40.2	228	2	hypothetical prote
19	43	40.2	229	2	ubiquitin 18 - sll
20	43	40.2	368	2	ubiquitin (clone 1)
21	43	40.2	380	2	hypothetical prote
22	43	40.2	380	2	polyubiquitin 5 (c
23	43	40.2	381	2	polyubiquitin 5 (c
24	43	40.2	532	2	polyubiquitin 7 (c
25	43	40.2	1810	1	tenascin precursor
26	42	39.3	204	2	probable peptide c
27	42	39.3	330	2	peptide chain rele
28	42	39.3	427	2	stress-activated p
29	42	39.3	672	2	penicillin-binding

30	42	39.3	1742	2	hypothetical prote
31	42	39.3	1959	2	hypothetical prote
32	41.5	38.8	376	1	chorismate synthas
33	41.5	38.8	500	2	4-aminobutyrate tr
34	41	38.3	166	2	translation releas
35	41	38.3	166	2	translation releas
36	41	38.3	166	2	probable peptide c
37	41	38.3	204	2	probable peptide c
38	41	38.3	260	2	protein F20B17.2
39	41	38.3	269	1	protein F20B17.2
40	41	38.3	313	2	protein B0495.8
41	41	38.3	333	2	alcohol dehydrogen
42	41	38.3	386	1	yeast protein - Bac
43	41	38.3	399	2	queuine tRNA-ribos
44	41	38.3	550	2	hypothetical prote
45	41	38.3	683	2	repressor protein

ALIGNMENTS

RESULT 1
OLECB
heat-labile enterotoxin chain B precursor - Escherichia coli
C:Species: Escherichia coli
C>Date: 29-Jun-1981 #sequence, revision 29-Jun-1981 #text, change 18-Jun-1999
C:Accession: A01820; B26946; I41194; I41287; I67644; A61475
R:Dallas, W.S.; Falkow, S.
Nature 288, 499-501, 1980
A>Title: Amino acid sequence homology between cholera toxin and Escherichia coli heat
A:Reference number: A01820; MVID:81074965
A:Accession: A01820
A:Molecule type: mRNA
A:Residues: 1-124 <DAL>
R:Yamamoto, T.; Gojohori, T.; Yokota, T.
J. Bacteriol. 169, 1352-1357, 1987
A>Title: Evolutionary origin of pathogenic determinants in enterotoxigenic Escherichia coli
A:Reference number: A26946; MVID:87137303
A:Accession: B26946
A:Molecule type: DNA
A:Residues: 1-27, 'E', '29-63, 'K', '65-124 <YAM>
A:Cross-references: EMBL:M5363; NID:g148335; PIDN:AAA24792.1; PID:g148336
R:Leong, J.; Vinal, A.C.; Dallas, W.S.
Infect. Immun. 48, 73-77, 1985
A>Title: Nucleotide sequence comparison between heat-labile toxin B-subunit cistrons
A:Reference number: I41194; MVID:85166481
A:Accession: I41194
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-5, 'F', '7-17, 'C', '19-24, 'S', '26-27, 'E', '29-33, 'H', '35-63, 'K', '65-66, 'A', '68-122
A:Cross-references: GB:M17874; NID:g145830; PIDN:AAA98064.1; PID:g145831
A:Experimental source: Plasmid ENR-R PCG86
R:Ibrahim, I.; Gentz, R.
J. Biol. Chem. 262, 10189-10194, 1987
A>Title: A functional interaction between the signal peptide and the translation appa
ratus.
A:Reference number: I41287; MVID:87280041
A:Accession: I41287
A>Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-22 <RE2>
A:Cross-references: GB:M17101; NID:g146375; PIDN:AAA23973.1; PID:g146376
R:Imoue, T.; Tsuji, T.; Koto, M.; Imamura, S.; Miyama, A.
FEBS Microbiol. Lett. 108, 157-161, 1993
A>Title: Amino acid sequence of heat-labile enterotoxin from chicken enterotoxigenic
A:Reference number: I53542; MVID:93252225
A:Accession: I67644
A>Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-17, 'C', '19, 'Y', '21-24, 'S', '26-27, 'E', '29-63, 'K', '65-66, 'A', '68-122, 'E', '124 <R
A:Cross-references: GB:S60731; NID:g408994; PIDN:AAC60441.1; PID:g408996
R:Tsuji, T.; Iida, T.; Honda, T.; Mawatani, T.; Nagahama, M.; Sakurai, J.; Wada, K.;
Microb. Pathog. 2, 381-390, 1987

```

A>Title: A unique amino acid sequence of the B subunit of a heat-labile enterotoxin iso
A:Reference number: A61475; MUID:89180953
A:Accession: A61475
A:Molecule type: Protein
A:Residues: 22-24,'S',26-27,'E',29-63,'K',65-66,'A',68-95,'A',97-122,'E',124 <TSU>
A:Experimental source: strain 240-3
C:Complex: the heat-labile enterotoxin molecule contains one A chain and five or six B c
C:Function:
A:Description: the biological activity of the toxin is produced by the A chain, which ac
C:Superfamily: cholera enterotoxin beta chain
C:Keywords: enterotoxin
F:1-21/Domain: signal sequence #status predicted <SIG>
F:22-124/Product: heat-labile enterotoxin chain B #status predicted <MAT>
F:30-107/Disulfide bonds: #status predicted

Query Match          100.0%; Score 107; DB 1; Length 124;
Best Local Similarity 100.0%; Pred. No. 5,9e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 GETFEVPEVPGSOHIDSQKAI 21
        |||||
Db       66 GETFEVPEVPGSOHIDSQKAI 86

RESULT      2
XVVCB
cholera enterotoxin, B chain precursor VC1456 [validated] - Vibrio cholerae (strain N165
N:Alternate names: enterotoxin beta chain
C:Species: Vibrio cholerae
C:Date: 24-Apr-1984 #sequence revision 01-Sep-2000 #text change 02-Feb-2001
C:Accession: S14624; S39238; S39241; H82196; JCI1078; S17666; PC1010; A05130; A01819; A3B
R:Dams, E.; de Wolf, M.; Dierick, W.
Submitted to the EMBL Data Library, March 1991
A:Description: Correction of the cholera toxin nucleotide sequence of the Vibrio cholera
A:Reference number: S14623
A:Accession: S14624
A:Molecule type: DNA
A:Residues: 1-124 <DM>
A:Cross-references: EMBL:X58786; NID:948420; PIDN:CAA41593.1; PID:948422
A:Experimental source: strain 2125
R:Riebens, M.; Holmgren, J.
Submitted to the EMBL Data Library, November 1993
A:Description: Structure and arrangement of the Cholera toxin genes in vibrio Cholera O1
A:Reference number: S39238
A:Accession: S39238
A:Molecule type: DNA
A:Residues: 1-124 <LEB>
A:Cross-references: EMBL:X76390; NID:9433856; PIDN:CAA53973.1; PID:9433857
A:Accession: S39241
A:Molecule type: DNA
A:Residues: 1-124 <LEW>
A:Cross-references: EMBL:X76391; NID:9433859; PIDN:CAA53976.1; PID:9433861
R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwin, M.L.; Dodson, R.J.;
Chadson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, H
L.; R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A>Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A:Reference number: A82035; MUID:20406833
A:Accession: H82196
A>Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-124 <HEI>
A:Cross-references: GB:AE004224; GB:AE003852; NID:96655952; PIDN:AAF94613.1; GSPDB:GN000
A:Experimental source: serogroup O1, strain N16561; biotype El Tor
R:Shu, C.H.; Cao, C.; Zhang, J.S.; Ma, Q.Q.
Chinese Biochem. J. 9, 395-399, 1993
A>Title: Nucleotide sequence analysis of the gene encoding the classical biotype V. chol
A:Reference number: JCI078
A:Accession: JCI078
A:Molecule type: DNA
A:Residues: 1-20,'Q',22-31,'Q',33-38,'H',40-49,'G',51-67,'T',69-124 <SHI>
A:Experimental source: classical biotype strain 569B

```

R:Dams, E.; de Wolf, M.; Dierick, W.
Biochim. Biophys. Acta 1090, 139-141, 1991
A>Title: Nucleotide sequence analysis of the CT operon of the Vibrio cholerae classical
A:Reference number: S17665; PMID:91355224
A:Accession: S17666
A:Molecule type: DNA
A:Residues: 1-38, 'H','40'-67, 'T','69'-124 <DAZ>
A:Cross-references: EMBL:X5785; NID:g4888; PIDN:CAAA1591.1; PID:g48890
R:Ma, Q.J.; Liu, C.X.; Xiong, L.S.; Yu, X.Q.
Sci. Sin. B Chem. Biol. Agric. Med. Earth Sci. 34, 274-280, 1991
A>Title: B subunit of cholera toxin produced in Escherichia coli.
A:Reference number: PC1010
A:Accession: PC1010
A:Molecule type: protein
A:Residues: 22-38, 'H','40'-41 <MAQ>
R:Meakins, J.J.; Swartz, D.J.; Pearson, G.D.N.; Harford, N.; Groyne, F.; de Wilde,
Nature 306, 551-557, 1983
A:Reference number: A93320; PMID:84068199
A:Accession: A05130
A:Molecule type: DNA
A:Residues: 1-32, 'S','34'-74, 'S','76'-124 <MEK>
A:Cross-references: GB:X00171; NID:g48347; PIDN:CAA24996.1; PID:g758351
J.Kurosky, A.; Markel, D.E.; Peterson, J.W.
J. Biol. Chem. 255, 7257-7264, 1977
A>Title: Covalent structure of the beta chain of cholera enterotoxin.
A:Reference number: A01819; PMID:78005537
A:Accession: A01819
A:Molecule type: protein
A:Residues: 22-38, 'H','40'-42, 'N','44'-67, 'T','69'-90, 'N','92'-124 <UR>
J.Lai, C.Y.
J. Biol. Chem. 252, 7249-7256, 1977
A>Title: Determination of the primary structure of cholera toxin B subunit.
A:Reference number: A38033; PMID:78005536
A:Accession: A38033
A:Molecule type: protein
A:Residues: 22-38, 'H','40'-42, 'N','44'-67, 'T','69','E','71'-90, 'N','92'-124 <LAJ>
A>Note: The difference at residue 70 made during preparation
R:Nakashima, Y.; Napierkowski, P.; Schafer, D.E.; Konigsberg, W.H.
FEBS Lett. 68, 275-278, 1976
A>Title: Primary structure of the B subunit of cholera enterotoxin.
A:Reference number: A38034; PMID:77026365
A:Accession: A38034
A:Molecule type: protein
A:Residues: 22-38, 'H','40'-67, 'T','69','E','71','GS','74'-75, 'VE','78'-86, 'Q','88'-99, 'Q','101'-103
R:Takaoka, T.; Matanabe, H.; Shimonishi, Y.
Eur. J. Biochem. 146, 503-508, 1985
A>Title: Facile identification of protein sequences by mass spectrometry.
A:Reference number: A21910; PMID:85126976
A:Accession: A21910
A:Molecule type: protein
A:Residues: 22-38, 'H','40'-42, 'N','44'-67, 'T','69'-90, 'N','92'-124 <PAK>
A:Experimental source: biotype Inaba 569B
A>Note: Asp-65 was partially deaminated to Asp
C:Comment: The authors translated the codon TCA for residue 33 as Tyr.
C:Genetics:
A:Gene: WC1456
A:Map position: 1
A:Complex: the cholera enterotoxin molecule contains three kinds of chains: an alpha
chain noncovalently with the subunit B, an aggregate of five beta chains
C:Function:
A:Description: involved in binding of the toxin to cell membranes
C:Keywords: enterotoxin; toxin
F:1-21/Domain: signal sequence #status predicted <SIG>
F:22-124/Product: cholera enterotoxin chain B #status predicted <MAT>
F:30-107/Disulfide bonds: #status experimental

Query Match	88.8%;	Score 95;	DB 1;	Length 124;
Best Local Similarity	90.5%;	Pred. No. 5.2e-08;		
Matches 19;	Conservative 0;	Mismatches 2;	Indels 0;	Gaps 0;
QY	1	GETFOVEVPSQIHDSQKKAI	21	

Db 66 GAIFOVEVPGSQHIDSOKKAI 86

RESULT 3

A86457

C:Species: Arabidopsis thaliana (mouse-ear cress)

C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 02-Nov-2001

C:Accession: A86457

R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,

Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Creasy, T.H.; Dewar, K.;

ansen, N.F.; Hughes, B.; Huizar, L.

Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.

C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maitl, R.; Marzita

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shim, P.; Southwick, A.M.; Sun, H.; Tallo

ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A>Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A:Reference number: A86141; MUID:21016719

A:Accession: A86457

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-255 <STO>

A:Cross-references: GB:AE005172; NID:g10645493; PIDN:AAG21605.1; GSPDB:GN00141

C:Genetics:

A:Map position: 1

Query Match 45.8%; Score 49; DB 2; Length 255;
Best Local Similarity 45.0%; Pred. No. 3.1;
Matches 9; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

OY 2 ETFFVEVPGSQHIDSOKKAI 21

Db 101 ETFFVSGGCGHNRKRSASV 120

RESULT 4

T04150

RAD23 protein homolog - rice

C:Species: Oryza sativa (rice)

C>Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 11-May-2000

C:Accession: T04150

R:Schultz, T.F.; Quatrano, R.S.

Plant Mol. Biol. 34, 557-562, 1997

A>Title: Characterization and expression of a rice RAD23 gene.

A:Reference number: Z08695; MUID:97369378

A:Accession: T04150

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-392 <SCH>

A:Cross-references: EMBL:U63530; NID:g1488296; PIDN:AAB65641.1; PID:g1488297

A:Experimental source: cv. Nipponbare

C:Genetics:

A:Gene: RAD23

C:Superfamily: unassigned ubiquitin-related proteins; ubiquitin homology

Query Match 43.9%; Score 47; DB 2; Length 392;
Best Local Similarity 42.9%; Pred. No. 10;
Matches 9; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

OY 1 GTFQVEVPGSQHIDSOKKAI 21

Db 10 GSTFOIEVDSQKVDYKRIT 30

RESULT 5

C96580

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Species: Arabidopsis thaliana (mouse-ear cress)

C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001

C:Accession: C96580

R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,

Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar,

ansen, N.F.; Hughes, B.; Huizar, L.

Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim,

C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maitl, R.; Marzita

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shim, P.; Southwick, A.M.; Sun, H.; Tallo

ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A>Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A:Reference number: A86141; MUID:21016719

A:Accession: C96580

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-91 <STO>

A:Cross-references: GB:AE005173; NID:g4587538; PIDN:AAD25769.1; GSPDB:GN00141

C:Genetics:

A:Gene: F15I1.6

A:Map position: 1

Query Match 43.0%; Score 46; DB 2; Length 91;
Best Local Similarity 47.6%; Pred. No. 3.2;
Matches 10; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

OY 1 GTFQVEVPGSQHIDSOKKAI 21

Db 58 GTFNLEVKSGSEITQVKNMI 78

RESULT 6

C86400

hypothetical protein T17H3.2 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Dec-2001

C:Accession: C86400

R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,

Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar,

ansen, N.F.; Hughes, B.; Huizar, L.

Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim,

C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maitl, R.; Marzita

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shim, P.; Southwick, A.M.; Sun, H.; Tallo

ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A>Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A:Reference number: A86141; MUID:21016719

A:Accession: C86400

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-574 <STO>

A:Cross-references: GB:AE005172; NID:g5668763; PIDN:AAD45990.1; GSPDB:GN00141

C:Genetics:

A:Map position: 1

Query Match 43.0%; Score 46; DB 2; Length 574;
Best Local Similarity 58.8%; Pred. No. 23;
Matches 10; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

OY 3 TFQVEVPGSQHIDSOKK 19

Db 19 TFFVDPSSQHIEVKK 35

RESULT 7

T49189

kinesin heavy chain-like protein - Arabidopsis thaliana

N:Alternate names: protein MAA21.110

C:Species: Arabidopsis thaliana (mouse-ear cress)

C>Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 02-Jun-2000

Db 541 QTFWADAPGDAKIDALRKA 559

RESULT 12

VFHJH

genome polypeptide 1b - murine hepatitis virus (strain JHM)

N:Alternate names: 1b protein

C:Species: murine hepatitis virus, MHV

C>Date: 31-Dec-1992 #sequence_revision 31-Dec-1992 #text_change 11-Jun-1999

C:Accession: B36815

R:Lee, H.T.; Shieh, C.K.; Gorbelenya, A.E.; Koonin, E.V.; La Monica, N.; Tuler, J.; Bagc

submitted to Genbank, February 1991

A:Description: The complete sequence (22 kilobases) of murine coronavirus gene 1 encodin

A:Reference number: A36815

A:Accession: B36815

A:Molecule type: genomic RNA

A:Residues: 1-2731 <LEE>

C:Cross-references: GB:M55148; NID:g331851; PIDN:AAA6458.1; PID:g331853

R:Lee, H.T.; Shieh, C.K.; Gorbelenya, A.E.; Koonin, E.V.; La Monica, N.; Tuler, J.; Bagc

Virology 180, 567-582, 1991

A:Title: The complete sequence (22 kilobases) of murine coronavirus gene 1 encoding the

A:Reference number: A38547; MUID:91111976

A:Contents: annotation

A>Note: neither nucleotide nor complete amino acid sequence is given

C:Comment: This protein may be translated as a 1a-1b polyprotein by a ribosomal frameshi

C:Genetics:

A:Gene: 1b

C:Superfamily: Infectious bronchitis virus RNA-directed RNA polymerase

C:Keywords: glycoprotein; nucleotidyltransferase; RNA biosynthesis

F:269,504,785,1164,1287,1524,1842,2196,2575,2630,2645,2665/Binding site: carbohydrate (A

Query Match 41.1%; Score 44; DB 1; Length 2731;

Best Local Similarity 69.2%; Pred. No. 2,5e+02;

Matches 9; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 2 ETEFOVEVPGSOHI 14

Db 1198 ETEFOVNPVNOHI 1210

RESULT 13

C84726

Probable receptor-like protein kinase [Imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C>Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001

C:Accession: C84726

R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;

M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.

euss, D.; Niemann, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J

Nature 402, 761-768, 1999

A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.

A:Reference number: A84420; MUID:20083487

A:Accession: C84726

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-641 <STO>

A:Cross-references: GB:AE002093; NID:g4887748; PIDN:AAD32284.1; GSPDB:GN00139

C:Genetics:

A:Gene: AT2g31880

A:Map position: 2

Query Match

Best Local Similarity 40.7%; Score 43.5; DB 2; Length 641;

Matches 11; Conservative 3; Mismatches 7; Indels 1; Gaps 1;

OY 1 GETFOVEVPGSQ-HIDSOKKAI 21

Db 359 GEVFKAEPLPGSGKTIAMVKVI 380

RESULT 14

UQDOR

ubiquitin / ribosomal protein CEP52 - slime mold (Dictyostelium discoideum)

N:Alternate names: ubiquitin fusion protein

C:Species: Dictyostelium discoideum

C>Date: 31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change 22-Jun-1999

C:Accession: S00357; A25863

R:Mueller-Taubenberger, A.; Westphal, M.; Jaeger, E.; Noegel, A.; Gerisch, G.

FEBS Lett. 229, 273-278, 1988

A:Title: Complete cDNA sequence of a Dictyostelium ubiquitin with a carboxy-terminal

A:Reference number: S00357; MUID:88152253

A:Accession: S00357

A:Molecule type: mRNA

A:Residues: 1-128 <MUE>

A:Cross-references: EMBL:X07210; NID:g7381; PIDN:CAA30183.1; PID:g7382

A:Experimental source: strain AX2-214

C:Genetics:

A:Gene: DUB1

C:Superfamily: ubiquitin/ribosomal protein CEP52; ribosomal protein CEP52 homology; u

C:Keywords: DNA binding; protein biosynthesis; protein degradation; ribosome; zinc fi

F:1-76/Product: ubiquitin #status predicted <UBI>

F:77-128/Product: ribosomal protein CEP52 #status predicted <CPH>

F:95-114/Region: zinc finger CCCC motif

F:121-128/Region: nuclear location signal

Query Match 40.2%; Score 43; DB 1; Length 128;

Best Local Similarity 42.9%; Pred. No. 14;

Matches 9; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

OY 1 GETFOVEVPGSOHIHDSOKKAI 21

Db 10 GKTITLVEGSDNIENYKAKI 30

RESULT 15

UQDOR7

ubiquitin / ribosomal protein S27a - slime mold (Dictyostelium discoideum)

C:Species: Dictyostelium discoideum

C>Date: 30-Jun-1991 #sequence_revision 30-Jun-1991 #text_change 21-Jul-2000

C:Accession: E34080

R:Omachi, T.; Giorda, R.; Shaw, D.R.; Ennis, H.L.

Biochemistry 28, 5226-5231, 1989

A:Title: Molecular organization of developmentally regulated Dictyostelium discoideum

A:Reference number: A34080; MUID:89352609

A:Accession: E34080

A:Molecule type: mRNA

A:Residues: 1-154 <OHM>

A:Cross-references: GB:J02858; GB:J02858; NID:g167940; PIDN:AAA33264.1; PID:g167941

C:Superfamily: ubiquitin / rat ribosomal protein S27a; ribosomal protein S27a homolog

C:Keywords: protein biosynthesis; protein degradation; ribosome

F:1-76/Product: ubiquitin #status predicted <MAT1>

F:77-154/Product: ribosomal protein S27a #status predicted <MAT2>

F:102-152/Domain: ribosomal protein S27a homology <RIB>

Query Match 40.2%; Score 43; DB 1; Length 154;

Best Local Similarity 42.9%; Pred. No. 17;

Matches 9; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

OY 1 GETFOVEVPGSOHIHDSOKKAI 21

Db 10 GKTITLVEGSDNIENYKAKI 30

Search completed: July 3, 2002, 08:56:34
Job time: 234 sec

Gencore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 3, 2002, 09:04:43 ; Search time 21.51 Seconds

(Without alignments)
37.802 Million cell updates/sec

Title: US-09-786-648-5

Perfect score: 107

Sequence: 1 GETFOVEYPGSQHIDSQKAI 21

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

105224

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	107	100.0	124	1 ELBP_ECOLI
2	101	94.4	124	1 ELBP_ECOLI
3	95	88.8	124	1 CHMP_VIBCH
4	45	42.1	2733	1 RRPB_CVMA5
5	44	41.1	263	1 YC43_ODOSI
6	44	41.1	467	1 IREF_HUMAN
7	44	41.1	467	1 IREF_MOUSE
8	44	41.1	2731	1 RRPB_CVMA5
9	43	40.2	76	1 UBIO_DICDI
10	43	40.2	1808	1 TENA_CHICK
11	42	39.3	855	1 ST14_HUMAN
12	41.5	38.8	376	1 ARQC_YEAST
13	41.5	38.8	500	1 GABP_HUMAN
14	41	38.3	141	1 RFB_ECOLI
15	41	38.3	313	1 YP68_CAEEL
16	41	38.3	385	1 TGT_ZYMO
17	41	38.3	386	1 YVAN_BACSU
18	41	38.3	567	1 SYP_STRCO
19	41	38.3	683	1 RPC_BPHC
20	41	38.3	1702	1 DPOL_THELI
21	41	38.3	1829	1 DPOL_THEST
22	40.5	37.9	862	1 PMS2_HUMAN
23	40	37.4	76	1 UBIO_NEUCR
24	40	37.4	148	1 IF5A_AERPE
25	40	37.4	308	1 HST4_CAEEL
26	40	37.4	289	1 CLT2_RAT
27	40	37.4	355	1 YRY1_CAEEL
28	40	37.4	395	1 ADH3_ENTHI
29	40	37.4	523	1 DBP3_YEAST
30	40	37.4	773	1 DPOL_THESG
31	40	37.4	775	1 DPOL_THES9
32	40	37.4	1048	1 CPXB_BACME
33	40	37.4	1523	1 DPOL_THPEM

ALIGNMENTS

RESULT	ID	ELBP_ECOLI	STANDARD	PRT	124 AA.
AC	P32890	P13768	P01557		
DT	21-JUL-1986	(Rel. 01, Created)			
DT	01-NOV-1995	(Rel. 32, Last sequence update)			
DT	15-JUL-1998	(Rel. 36, Last annotation update)			
DE	Heat-labile enterotoxin B chain precursor (Lr-B, porcine) (LTP-B).				
GN	ELTB OR LTPB.				
OS	Escherichia coli.				
OC	Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;				
OC	Escherichia.				
OX	NCBI_TaxID=562;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN-ISOLATE P307;				
RX	MEDLINE=81074965; PubMed=7003397;				
RA	Dallas W.S., Falkow S.;				
RT	"Amino acid sequence homology between cholera toxin and Escherichia				
RT	coli heat-labile toxin.";				
RL	Nature 288:499-501(1980).				
RN	[2]				
RP	REVISIONS TO 28 AND 64.				
RC	STRAIN-ISOLATE P307;				
RX	MEDLINE=85156481; PubMed=3884513;				
RA	Jeong J., Vinal A.C., Dallas W.S.;				
RT	"Evolutionary origin of pathogenic determinants in enterotoxigenic				
RT	Escherichia coli and Vibrio cholerae O1.";				
RL	J. Bacteriol. 169:1352-1357(1987).				
RN	[3]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN-ISOLATE PCG86;				
RX	MEDLINE=87137303; PubMed=3546273;				
RA	Yamamoto T., Gojobori T., Yokota T.;				
RT	"A functional interaction between the signal peptide and the				
RT	translation apparatus is detected by the use of a single point				
RT	mutation which blocks translocation across mammalian endoplasmic				
RT	reticulum.";				
RL	J. Biol. Chem. 262:10189-10194(1987).				
RN	[5]				
RP	X-RAY CRYSTALLOGRAPHY (1.95 ANGSTROMS).				
RX	MEDLINE=93240541; PubMed=8478941;				
RA	Sixma T.K., van Zanten B.A.M., Dauter Z., Hol W.G.J.;				
RT	"Refined structure of Escherichia coli heat-labile enterotoxin, a				
RT	close relative of cholera toxin.";				
RL	J. Mol. Biol. 230:890-918(1993).				
RN	[6]				
RP	X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS).				
RX	MEDLINE=91238966; PubMed=2034287;				

34	40	37.4	1668	1	DPOL_THEHY	09h05 thermococcus
35	40	37.4	1671	1	DPOL_PYRKO	P77933 pyrococcus
36	40	37.4	1699	1	DPOL_THES8	Q9h84 thermococcus
37	39.5	36.9	1090	1	PULA_KLEPN	P07206 klebsiella
38	39	36.4	76	1	UBIO_YEAST	P04838 saccharomyc
39	39	36.4	271	1	CEAM_ECOLI	P05820 escherichia
40	39	36.4	320	1	P7504_MYCPN	P75049 mycoplasma
41	39	36.4	445	1	RPM_HUMAN	O75570 homo sapien
42	39	36.4	456	1	SR54_THESAC	Q9hct0 thermoplasm
43	39	36.4	459	1	IL7R_MOUSE	P16872 mus musculu
44	39	36.4	491	1	CD5_RAT	P51882 ratus norv
45	39	36.4	560	1	INRL_SHEEP	O28589 ovis aries

RA Sijma T.K., Pronk S.E., Kalk K.H., Wartna E.S., van Zanten B.A.M.,
 RA Wittholt B., Hol W.G.J.;
 RT "Crystal structure of a cholera toxin-related heat-labile enterotoxin
 from *E. coli*.";
 RL Nature 351:371-377(1991).
 RN [7]
 RP DISCUSSION OF SEQUENCE.
 RX MEDLINE=95349400; PubMed=7623669;
 RA Domenighini M., Pizzi M., Jobling M.G., Holmes R.K., Rappuoli R.;
 RT "Identification of errors among database sequence entries and
 RT comparison of correct amino acid sequences for the heat-labile
 RT enterotoxin of *Escherichia coli* and *Vibrio cholerae*.";
 RL Mol. Microbiol. 15:1165-1167(1995).
 CC -I- FUNCTION: THE BIOLOGICAL ACTIVITY OF THE TOXIN IS PRODUCED BY
 CC THE A CHAIN, WHICH ACTIVATES INTRACELLULAR ADENYL CYCLASE.
 CC -I- SUBUNIT: HETEROHEXAMER OF ONE A CHAIN AND OF FIVE B CHAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL, M17873; AAA98065.1; -;
 DR EMBL, M15363; AAA24792.1; -;
 DR EMBL, M17101; AAA23973.1; -;
 DR PIR, A01820; QLECB.
 DR PIR, B26946; QLECB.
 DR PDB, 1LTA; 31-JAN-94.
 DR PDB, 1LTB; 31-JAN-94.
 DR PDB, 1LTG; 15-SEP-95.
 DR PDB, 1LTI; 17-AUG-96.
 DR PDB, 1LTS; 31-JAN-94.
 DR PDB, 1LTT; 31-JAN-94.
 DR PDB, 1LTJ; 07-JUL-97.
 DR PDB, 1LT4; 16-JUN-97.
 DR PDB, 1LT5; 03-DEC-97.
 DR PDB, 1LT6; 03-DEC-97.
 DR PDB, 1LTL; 20-APR-95.
 DR InterPro: IPR001835; Enterotoxin_B.
 DR Pfam: PF01376; Enterotoxin_B.
 DR PRINTS: PR00772; ENTEROTOXINB.
 DR ProDom: PD012805; Enterotoxin_B; 1.
 KW Enterotoxin; Signal; 3D-structure.
 FT SIGNAL 1 21
 FT CHAIN 22 124 HEAT-LABILE ENTEROTOXIN B CHAIN.
 FT DISULFID 30 107
 FT HELIX 26 30
 FT TURN 31 32
 FT STRAND 36 43
 FT STRAND 47 51
 FT TURN 54 55
 FT STRAND 58 62
 FT TURN 64 65
 FT STRAND 68 71
 FT TURN 76 77
 FT HELIX 80 98
 FT TURN 99 100
 FT STRAND 103 109
 FT STRAND 115 123
 SO SEQUENCE 124 AA; 14133 MW; 6DB7DE58395EA70D CRC64;

Query Match 100.0%; Score 107; DB 1; Length 124;
 Best Local Similarity 100.0%; Pred. No. 3e-10;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 GETFOVEVPGSOHDSOKKAI 21
 ||||||||||||||||
 Db 66 GETFOVEVPGSOHDSOKKAI 86

RESULT 2
 EBLH_ECOLI STANDARD; PRT; 124 AA.
 ID EBLH_ECOLI
 AC P13811;
 DT 01-JAN-1990 (Rel. 13, Created)
 DT 01-JAN-1990 (Rel. 13, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Heat-labile enterotoxin B chain precursor (LT-B, human) (LTH-B).
 GN ELPB OR LTPB.
 OS *Escherichia coli*.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC *Escherichia*.
 OX NCBI_TaxID=562;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ISOLATE H74-114;
 RX MEDLINE=85156481; PubMed=3884513;
 RA Leong J., Vinal A.C., Dallas W.S.;
 RT "Nucleotide sequence comparison between heat-labile toxin B-subunit
 RT cistrons from *Escherichia coli* of human and porcine origin.";
 RL Infect. Immun. 48:73-77(1985).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ISOLATE H10407;
 RX MEDLINE=83114628; PubMed=6759877;
 RA Yamamoto T., Tamura T.A., Yokota T., Takano T.;
 RT "Overlapping genes in the heat-labile enterotoxin operon originating
 RT from *Escherichia coli* human strain.";
 RL Mol. Gen. Genet. 188:356-359(1982).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ISOLATE H10407;
 RX MEDLINE=93252225; PubMed=8486242;
 RA Inoue T., Tsuji T., Koto M., Imanura S., Miyama A.;
 RT "Amino acid sequence of heat-labile enterotoxin from chicken
 RT enterotoxigenic *Escherichia coli* is identical to that of human strain
 RT H 10407.";
 RL FEMS Microbiol. Lett. 108:157-161(1993).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ETEC LT 87;
 RA Germain Y., Desperrier J.M.;
 RL Submitted (Oct-1996) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP DISCUSSION OF SEQUENCE.
 RX MEDLINE=95349400; PubMed=7623669;
 RA Domenighini M., Pizzi M., Jobling M.G., Holmes R.K., Rappuoli R.;
 RT "Identification of errors among database sequence entries and
 RT comparison of correct amino acid sequences for the heat-labile
 RT enterotoxins of *Escherichia coli* and *Vibrio cholerae*.";
 RL Mol. Microbiol. 15:1165-1167(1995).
 RN [6]
 RP X-RAY CRYSTALLOGRAPHY (3.04 ANGSTROMS).
 RX MEDLINE=99185101; PubMed=10085117;
 RA Matkovic-Calogovic D., Loregian A., D'Acunzio M.R., Battistutta R.,
 RA Tossi A., Palu G., Zanotti G.;
 RT "Crystal structure of the B subunit of *Escherichia coli* heat-labile
 RT enterotoxin carrying peptides with anti-herpes simplex virus type 1
 RT activity.";
 RL J. Biol. Chem. 274:8764-8769(1999).
 CC -I- FUNCTION: THE BIOLOGICAL ACTIVITY OF THE TOXIN IS PRODUCED BY
 CC THE A CHAIN, WHICH ACTIVATES INTRACELLULAR ADENYL CYCLASE.
 CC -I- SUBUNIT: HETEROHEXAMER OF ONE A CHAIN AND OF FIVE B CHAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----

DR EMBL: M17874; AAA98064.1; -
 DR EMBL: J01646; AAB02982.1; -
 DR EMBL: S60731; AAC60441.1; -
 DR EMBL: X83966; CAA58800.1; -
 DR PDB: 1L7R; 23-MAR-99.
 DR InterPro: IPR001835; Enterotoxin_B.
 DR Pfam: PF01376; Enterotoxin_B; 1.
 DR PRINTS: PR00772; ENTEROTOXINB.
 DR ProDom: PD012805; Enterotoxin_B; 1.
 DR Enterotoxin; Signal; 3D-structure.
 FT SIGNAL 1 21
 FT CHAIN 22 124 HEAT-LABILE ENTEROTOXIN B CHAIN.
 FT DISULFID 30 107
 FT VARIANT 6 6 F -> C (IN ISOLATE H10407).
 FT VARIANT 20 20 H -> Y (IN ISOLATE H10407).
 FT VARIANT 34 34 H -> R (IN ISOLATE H10407).
 FT SEQUENCE 124 AA; 14027 MW; E9F7FC7B9D3BC47 CRC64;
 SQ
 Query Match 94.4%; Score 101; DB 1; Length 124;
 Best Local Similarity 95.2%; Pred. No. 2.7e-09;
 Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 OY 1 GETFOVEVPGSOHIDSKKAI 21
 DB 66 GATFOVEVPGSOHIDSKKAI 86
 RESULT 3
 ID CHTB_VIBCH STANDARD; PRT; 124 AA.
 AC P01556; Q9J002;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 13-AUG-1987 (Rel. 05, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Cholera enterotoxin, beta chain precursor.
 GN CTXB OR TOXB OR VC1456.
 OS Vibrio cholerae.
 OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
 OX NCBI_TaxID=66;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=84061784; PubMed=6315707;
 RA Lockman H., Kaper J.B.;
 RT "Nucleotide sequence analysis of the A2 and B subunits of Vibrio
 cholerae enterotoxin.";
 RL J. Biol. Chem. 258:13722-13726(1983).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-EL TOR 2125;
 RX MEDLINE=84068199; PubMed=6646234;
 RA Mekalanos J.J., Swartz D.J., Pearson G.D.N., Harford N., Groyne F.,
 de Wilde M.;
 RT "Cholera toxin genes: nucleotide sequence, deletion analysis and
 vaccine development.";
 RL Nature 306:551-557(1983).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-EL TOR 2125;
 RA Dams E., de Wolf M., Dierick W.;
 RL Submitted (May-1991) to the EMBL/Genbank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=4260B / SEROTYPE O139;
 RX MEDLINE=94237453; PubMed=8181723;
 RA Lebens M., Holmgren J.;
 RT "Structure and arrangement of the cholera toxin genes in Vibrio
 cholerae O139.";
 RL FEMS Microbiol. Lett. 117:197-202(1994).
 RN [5]
 RP SEQUENCE FROM N.A.
 RC STRAIN=1854 / O139-BENGAL;
 RA Yamamoto K., Do V.G.R.F., Xu M., Iida T., Miwatani T., Albert M.J.,

RA Honda T.;
 RL Submitted (May-1994) to the EMBL/Genbank/DBJ databases.
 RN [6]
 RP SEQUENCE FROM N.A.
 RC STRAIN-EL TOR N16961 / SEROTYPE O1;
 RX MEDLINE=20406833; PubMed=10952301;
 RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwin M.L.,
 Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,
 Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
 Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,
 McDonald L., Utterback T., Fleischmann R.D., Nielsen W.C., White O.,
 Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
 Fraser C.M.;
 RA "DNA sequence of both chromosomes of the cholera pathogen Vibrio
 cholerae.";
 RL Nature 406:477-483(2000).
 RN [7]
 RP SEQUENCE OF 22-124.
 RX MEDLINE=78005537; PubMed=903363;
 RA Kurosky A., Markel D.E., Peterson J.W.;
 RT "Covalent structure of the beta chain of cholera enterotoxin.";
 RL J. Biol. Chem. 252:7257-7264(1977).
 RN [8]
 RP SEQUENCE OF 22-124.
 RX MEDLINE=78005536; PubMed=903362;
 RA Lai C.-Y.;
 RT "Determination of the primary structure of cholera toxin B subunit.";
 RL J. Biol. Chem. 252:7249-7256(1977).
 RN [9]
 RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS).
 RX MEDLINE=94272319; PubMed=8003954;
 RA Merritt E.A., Sarfaty S., van den Akker F., L'Hoir C., Martini J.A.,
 Hol W.G.J.;
 RT "Crystal structure of cholera toxin B-pentamer bound to receptor GM1
 pentasaccharide.";
 RL Protein Sci. 3:166-175(1994).
 RN [10]
 RP X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS).
 RX MEDLINE=95387394; PubMed=7658472;
 RA Zhang R.G., Westbrook M.L., Westbrook E.M., Scott D.L., Orlowski Z.,
 Maulik P.R., Reed R.A., Shipley G.G.;
 RT "The 2.4 A crystal structure of cholera toxin B subunit pentamer:
 choleraenoid.";
 RL J. Mol. Biol. 251:550-562(1995).
 RN [11]
 RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).
 RC STRAIN-OGAWA 41 / CLASSICAL BIOTYPE;
 RX MEDLINE=97376625; PubMed=9232653;
 RA Merritt E.A., Sarfaty S., Jobling M.G., Chang T., Holmes R.K.,
 Hirst T.R., Hol W.G.J.;
 RT "Structural studies of receptor binding by cholera toxin mutants.";
 RL Protein Sci. 6:1516-1528(1997).
 CC -1- FUNCTION: THE BETA CHAIN AGGREGATE (B SUBUNIT) IS INVOLVED IN
 BINDING TO CELL MEMBRANES.
 CC -1- SUBUNIT: CONTAINS 3 KINDS OF CHAINS. AN ALPHA AND A GAMMA CHAIN
 (FROM THE SAME PRECURSOR MOLECULE), LINKED BY AN INTERCHAIN
 DISULFIDE BOND, ASSOCIATE NONCOVALENTLY WITH AN AGGREGATE OF 4 TO
 6 BETA CHAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: X00171; CAA24996.1; -
 DR EMBL: K01170; AAA27573.1; -
 DR EMBL: D30053; BAA06291.1; -
 DR EMBL: X58786; CAA41593.1; -
 DR EMBL: X76390; CAA53973.1; -
 DR EMBL: X76391; CAA53976.1; -

RX MEDLINE-91111976; PubMed-1846489;
RA Lee H.-J., Shieh C.-K., Gorbaleya A.E., Koonin E.V., la Monica N.,
RA Tuler J., Bagdzhardzhan A., Lai M.M.C.;
RT "The complete sequence (22 kilobases) of murine coronavirus gene 1
RT encoding the putative proteases and RNA polymerase.";
RL Virology 180:567-582(1991).
CC -1- FUNCTION: THE RNA DEPENDENT RNA POLYMERASE OF CORONAVIRUSES IS
CC A MULTIFUNCTIONAL PROTEIN. IT CONTAINS THE ACTIVITIES NECESSARY
CC FOR THE TRANSCRIPTION OF NEGATIVE STRANDED RNA, LEADER RNA,
CC SUBGENOMIC MRNAS AND PROGENY VIRION RNA.
CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate -> N diphosphate +
CC (RNA)(N).
CC -1- MISCELLANEOUS: THIS PROTEIN MAY BE TRANSLATED AS A 1A-1B
CC POLYPEPTIDE BY A RIBOSOMAL FRAMESHIFTING MECHANISM.
CC -1- SIMILARITY: TO THE AVIAN CORONAVIRUS IBV RNA POLYMERASE.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M55148; AAA46458.2; -;
DR PIR; B36815; VFTHJH.
KW Transferase; RNA-directed RNA polymerase; Helicase; ATP-binding.
FT DOMAIN 622 824 POLYMERASE.
FT DOMAIN 944 1014 CYS/HIS-RICH.
FT DOMAIN 1218 1506 HELICASE.
FT NP_BIND 1220 1227 ATP (BY SIMILARITY).
SQ SEQUENCE 2731 AA; 308852 MW; 99463066B176970 CRC64;

Query Match 41.1%; Score 44; DB 1; Length 2731;
Best Local Similarity 69.2%; Pred. No. 1e+02;
Matches 9; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 ETPOVEVPGSOHI 14
Db 1198 ETPONNPNVNGHI 1210

RESULT 9
UBIO_DICDI STANDARD; PRT; 76 AA.
AC P08618;
DT 01-AUG-1988 (Rel. 08, Created)
DT 01-AUG-1988 (Rel. 08, Last sequence update)
DT 01-NOV-1991 (Rel. 20, Last annotation update)
DE Ubiquitin.
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyostellida; Dictyostellium.
OX NCBI_TaxID=44689;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-88152253; PubMed-2831095;
RA Mueller-Taubenberger A., Westphal M., Jaeger E., Noegel A.,
RA Gerisch G.;
RT "Complete cDNA sequence of a Dictyostelium ubiquitin with a carboxy-
RT terminal tail and identification of the protein using an anti-peptide
RT antibody.";
RL FEBS Lett. 229:273-278(1988).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE-89352609; PubMed-2548604;
RA Omachi T., Giorda R., Shaw D.R., Ennis H.L.;
RT "Molecular organization of developmentally regulated Dictyostelium
RT discoideum ubiquitin cDNAs.";
RL Biochemistry 28:5226-5231(1989).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE-87257921; PubMed-3037345;

RA Giorda R., Ennis H.L.;
RT "Structure of two developmentally regulated Dictyostelium discoideum
RT ubiquitin genes.";
RL Mol. Cell. Biol. 7:2097-2103(1987).
RN [4]
RP SEQUENCE OF 13-76 FROM N.A.
RC STRAIN-A42;
RA Westphal M., Mueller-Taubenberger A., Noegel A., Gerisch G.;
RT "Transcript regulation and carboxyterminal extension of ubiquitin in
RT Dictyostelium discoideum.";
RL FEBS Lett. 209:92-96(1986).
CC -1- FUNCTION: INVOLVED IN THE ATP-DEPENDENT SELECTIVE DEGRADATION OF
CC CELLULAR PROTEINS, THE MAINTENANCE OF CHROMATIN STRUCTURE, THE
CC REGULATION OF GENE EXPRESSION, THE STRESS RESPONSE, AND RIBOSOME
CC BIOGENESIS.
CC -1- SUBCELLULAR LOCATION: Nuclear and cytoplasmic.
CC -1- MISCELLANEOUS: UBIQUITIN IS SYNTHESIZED AS A POLYUBIQUITIN
CC PRECURSOR WITH EXACT HEAD TO TAIL REPEATS. SOME UBIQUITIN GENES
CC CONTAIN A SINGLE COPY OF UBIQUITIN FUSED TO A RIBOSOMAL PROTEIN.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X07210; CAA30183.1; ALT_TERM.
DR EMBL; M19666; AAA33261.1; ALT_TERM.
DR EMBL; M19491; AAA33269.1; ALT_TERM.
DR EMBL; M19492; AAA33270.1; ALT_TERM.
DR EMBL; M23748; AAA33262.1; ALT_TERM.
DR EMBL; M23749; AAA33263.1; ALT_TERM.
DR EMBL; M23750; AAA33264.1; ALT_TERM.
DR EMBL; M23751; AAA33265.1; ALT_TERM.
DR EMBL; M23752; AAA33266.1; ALT_TERM.
DR EMBL; M23753; AAA33267.1; ALT_TERM.
DR EMBL; M23754; AAA33268.1; ALT_TERM.
DR EMBL; X04702; CAA28408.1; ALT_TERM.
DR PIR; S00357; UODOR.
DR PIR; E34080; UODOR.
DR PIR; A27806; A27806.
DR PIR; B27806; B27806.
DR PIR; A34080; A34080.
DR PIR; B34080; B34080.
DR PIR; C34080; C34080.
DR PIR; D34080; D34080.
DR HSSP; P02248; I0B1.
DR Dictydb; DD05001; ubqa.
DR InterPro; IPR000626; Ubiquitin.
DR Pfam; PF00240; ubiquitin; 1.
DR PRINTS; PR00348; UBIQUITIN.
DR SMART; SM00213; UBQ; 1.
DR PROSITE; PS00299; UBIQUITIN_1; 1.
DR PROSITE; PS00053; UBIQUITIN_2; 1.
KW Nuclear protein; Polypeptide.
FT SITE 48 48
FT BINDING 76 76
FT VARIANT 11 11
SQ SEQUENCE 76 AA; 8538 MW; 6427383968EA8A84 CRC64;

Query Match 40.2%; Score 43; DB 1; Length 76;
Best Local Similarity 42.9%; Pred. No. 3.2;
Matches 9; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

QY 1 GETPOVEVPGSOHIIDOKKAI 21
Db 10 KKTITLVEGSDNIENYAKAI 30

FT DISULFID 413 424 BY SIMILARITY.
 FT DISULFID 426 435 BY SIMILARITY.
 FT DISULFID 440 450 BY SIMILARITY.
 FT DISULFID 444 455 BY SIMILARITY.
 FT DISULFID 457 466 BY SIMILARITY.
 FT DISULFID 471 481 BY SIMILARITY.
 FT DISULFID 475 486 BY SIMILARITY.
 FT DISULFID 488 497 BY SIMILARITY.
 FT DISULFID 502 512 BY SIMILARITY.
 FT DISULFID 506 517 BY SIMILARITY.
 FT DISULFID 519 528 BY SIMILARITY.
 FT DISULFID 533 543 BY SIMILARITY.
 FT DISULFID 537 548 BY SIMILARITY.
 FT DISULFID 550 559 BY SIMILARITY.
 FT DISULFID 564 574 BY SIMILARITY.
 FT DISULFID 568 579 BY SIMILARITY.
 FT DISULFID 581 590 BY SIMILARITY.
 FT CARBOHYD 38 38 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 168 168 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 186 186 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 328 328 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 603 603 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 643 643 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 751 751 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 759 759 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1050 1050 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1090 1090 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1101 1101 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1112 1112 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1153 1153 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1183 1183 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1416 1416 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1736 1736 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1769 1769 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT VARSPLIC 1043 1234 MISSING (IN 200 KDA ISOFORM).
 FT VARSPLIC 1043 1315 MISSING (IN 190 KDA ISOFORM).
 FT CONFLICT 182 182 W -> R (IN REF. 2).
 FT CONFLICT 563 571 SCPNDCNNV -> PAPMTWTW (IN REF. 3).
 FT CONFLICT 598 598 E -> G (IN REF. 3).
 FT CONFLICT 838 838 T -> TEY (IN REF. 3).
 FT CONFLICT 886 886 T -> F (IN REF. 3).
 SQ SEQUENCE 1808 AA; 198858 MW; B924A06CF9E9D6DE CRC64;

Query Match 40.2%; Score 43; DB 1; Length 1808;
 Best Local Similarity 58.3%; Pred. No. 93;
 Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 2 ETRQVEVPSQH 13

Db 1087 EFWNITVPGQH 1098

RESULT 11
 ID ST14_HUMAN STANDARD; PRT; 855 AA.
 AC Q9Y5V6; Q9H3S0; Q9HCA3; Q9BS01; Q9HB36;
 DT 16-OCT-2001 (Rel. 40; Created)
 DT 16-OCT-2001 (Rel. 40; Last sequence update)
 DE Suppressor of tumorigenicity 14 (bc 3.4.21.-) (Matrilinease) (Membrane-type serine protease 1) (MT-SPI) (Proteinase) (Serine protease TADG-15)
 DE (Tumor associated differentially-expressed gene-15 protein).
 GN ST14 OR PRSS14 OR SMC19 OR TADG15.
 OS Homo sapiens (Human)
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99303581; PubMed=10373424;
 RA Lin C.Y., Anders J., Johnson M., Sang Q.A., Dickson R.B.;
 RT "Molecular cloning of cDNA for matrilysin, a matrix-degrading serine

RT protease with trypsin-like activity.";
 RL J. Biol. Chem. 274:18231-18236(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99432178; PubMed=10500122;
 RA Takeuchi T., Shuman M.A., Craik C.S.;
 RT "Reverse biochemistry: use of macromolecular protease inhibitors to
 RT dissect complex biological processes and identify a membrane-type
 RT serine protease in epithelial cancer and normal tissue.";
 RL Proc. Natl. Acad. Sci. U.S.A. 96:11054-11061(1999).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Prostate;
 RA Yamaguchi N., Mitsui S.;
 RT "Molecular cloning of a novel transmembrane serine protease expressed
 RT in human prostate.";
 RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RA Tanimoto H., Underwood L.J., Wang Y., Shigemasa K., Parmley T.H.,
 RA O'Brien T.J.;
 RT Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP SEQUENCE OF 327-855 FROM N.A.
 RC TISSUE=Muscle;
 RA Strausberg R.;
 RT Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
 RN [6]
 RP SEQUENCE OF 340-664 FROM N.A.
 RA Cao J., Fan W., Zheng S.;
 RT "Genomic analysis of a novel human serine protease SMC19.";
 RN Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
 RP CHARACTERIZATION.
 RC TISSUE=Milk;
 RX MEDLINE=99303582; PubMed=10373425;
 RA Lin C.Y., Anders J., Johnson M., Dickson R.B.;
 RT "Purification and characterization of a complex containing matrilysin
 RT and a kunitz-type serine protease inhibitor from human milk.";
 RL J. Biol. Chem. 274:18237-18242(1999).
 CC -1- FUNCTION: DEGRADERS EXTRACELLULAR MATRIX. PROPOSED TO PLAY A ROLE
 CC IN BREAST CANCER INVASION AND METASTASIS. EXHIBITS TRYPSIN-LIKE
 CC ACTIVITY AS DEFINED BY CLEAVAGE OF SYNTHETIC SUBSTRATES WITH ARG
 CC OR LYS AS THE P1 SITE.
 CC -1- SUBCELLULAR LOCATION: Type II membrane protein (Probable).
 CC -1- SIMILARITY: CONTAINS 4 LDL-RECEPTOR CLASS A DOMAINS.
 CC -1- SIMILARITY: CONTAINS 2 CUB DOMAINS.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
 CC TRYPSIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: AF118224; AADA2765.2; -;
 DR EMBL: AF133086; AAF00109.1; -;
 DR EMBL: AB030036; BAB20376.1; -;
 DR EMBL: AF057145; AAG15395.1; -;
 DR EMBL: BC005826; AAH05826.1; -;
 DR EMBL: AF283256; AAG13949.1; -;
 DR HSSP: P00763; IDPO.
 DR MEROPS: S01.302; -;
 DR InterPro: IPR000859; CUB.
 DR InterPro: IPR001314; Chymotrypsin.
 DR InterPro: IPR002172; LDL_recept_A.
 DR InterPro: IPR001254; Trypsin.
 DR Pfam: PF00431; CUB; 2.
 DR Pfam: PF00057; LDL_recept_a; 4.
 DR Pfam: PF00089; trypsin; 1.

DR PRINTS; PR00722; CHYMOTRYPSIN.
 DR PRINTS; PR00261; LDLRECEPTOR.
 DR SMART; SM00042; CUB; 2.
 DR SMART; SM00192; LDLA; 3.
 DR SMART; SM00020; TRYP_SPC; 1.
 DR PROSITE; PS01180; CUB; 2.
 DR PROSITE; PS01209; LDLRA_1; 2.
 DR PROSITE; PS50068; LDLRA_2; 4.
 DR PROSITE; PS50240; TRYPIN_DOM; 1.
 DR PROSITE; PS00134; TRYPIN_HIS; 1.
 DR PROSITE; PS00135; TRYPIN_SER; 1.
 KW Signal anchor; Glycoprotein; Hydrolase; Serine protease;
 KW Transmembrane; Repeat.
 FT TRANSMM 1 55
 FT DOMAIN 56 76
 FT FT
 FT DOMAIN 77 85
 FT DOMAIN 214 334
 FT DOMAIN 340 447
 FT DOMAIN 452 487
 FT DOMAIN 487 524
 FT DOMAIN 524 560
 FT DOMAIN 566 603
 FT DOMAIN 615 854
 FT ACT_SITE 656 656
 FT ACT_SITE 711 711
 FT ACT_SITE 711 805
 FT CARBOHYD 109 109
 FT CARBOHYD 302 302
 FT CARBOHYD 485 485
 FT CARBOHYD 772 772
 FT CONFLICT 327 329
 FT CONFLICT 381 381
 FT CONFLICT 674 674
 SQ SEQUENCE 855 AA; 94769 MW; 26143132C01F99C9 CRC64;

Query Match 39.3%; Score 42; DB 1; Length 855;
 Best Local Similarity 50.0%; Pred. No. 61;
 Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

OY 3 TFOVEVPGSQHI 14
 I:||||:|:
 DB 367 TWNIEVPNNQHV 378

RESULT 12
 AROC_YEAST STANDARD; PRT; 376 AA.
 ID AROC_YEAST
 AC P28777;
 DT 01-DEC-1992 (Rel. 24, Created)
 DT 01-DEC-1992 (Rel. 24, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Chorismate synthase (EC 4.6.1.4) (5-enolpyruvylshikimate-3-phosphate
 DE phosphoylase).
 GN ARO2 OR YGL148V.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S288C;
 RX MEDLINE=92114793; PubMed=1837329;
 RA Jones D.G.U., Reusser U., Braus G.H.;
 RT "Molecular cloning, characterization and analysis of the regulation
 RT of the ARO2 gene, encoding chorismate synthase, of Saccharomyces
 RT cerevisiae.";
 RL Mol. Microbiol. 5:2143-2152(1991).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S288C / FY1769;
 RX MEDLINE=97197983; PubMed=9046099;

RA Voet M., Defoor E., Verhasselt P., Riles L., Robben J., Volckaert G.;
 RT "The sequence of a nearly unclonable 22.8 kb segment on the left arm
 RT chromosome VII from Saccharomyces cerevisiae reveals ARO2, RPL9A,
 RT TIF1, MRF1 genes and six new open reading frames";
 RL Yeast 13:177-182(1997).
 CC -1- CATALYTIC ACTIVITY: 5-O-(1-carboxyvinyl)-3-phosphoshikimate =
 CC chorismate + phosphate.
 CC -1- COFACTOR: REDUCED FLAVIN.
 CC -1- PATHWAY: SEVENTH STEP IN THE BIOSYNTHESIS OF CHORISMATE WITHIN
 CC THE BIOSYNTHESIS OF AROMATIC AMINO ACIDS (THE SHIKIMATE PATHWAY).
 CC -1- SUBUNIT: HOMOTETRAMER (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE CHORISMATE SYNTHASE FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC
 CC EMBL; X60190; CAA42745.1; -;
 CC EMBL; X99960; CAA68214.1; -;
 CC EMBL; Z72670; CAA96860.1; -;
 CC PIR; S17246; S17246.
 CC SGD; S0003116; ARO2.
 CC InterPro; IPR000453; Chorismate_synth.
 CC Pfam; PF01264; Chorismate_synth; 1.
 CC ProDom; PD002941; Chorismate_synth; 1.
 CC PROSITE; PS00787; CHORISMATE SYNTHASE 1; 1.
 CC PROSITE; PS00788; CHORISMATE SYNTHASE 2; 1.
 CC PROSITE; PS00789; CHORISMATE SYNTHASE 3; 1.
 KW Lyase; Aromatic amino acid biosynthesis.
 SQ SEQUENCE 376 AA; 40838 MW; AF3AF65605B91E8E CRC64;

Query Match 38.8%; Score 41.5; DB 1; Length 376;
 Best Local Similarity 62.5%; Pred. No. 31;
 Matches 10; Conservative 1; Mismatches 4; Indels 1; Gaps 1;

OY 1 GETPQ-VEVPGSQHI 15
 I:||||:|:
 DB 258 GSGFGVSPGSKND 273

RESULT 13
 GABT_HUMAN STANDARD; PRT; 500 AA.
 ID GABT_HUMAN
 AC P80404;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE 4-aminobutyrate aminotransferase, mitochondrial precursor
 DE (EC 2.6.1.19) (Gamma-amino-N-butyrate transaminase) (GABA
 DE transaminase) (GABA aminotransferase) (GABA-AT).
 GN ABAT OR GABAT.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=95237607; PubMed=7721088;
 RA Osei Y.D., Churchich J.E.;
 RT "Screening and sequence determination of a cDNA encoding the human
 RT brain 4-aminobutyrate aminotransferase.";
 RL Gene 155:185-187(1995).
 RN [2]
 RP SEQUENCE OF 368-465 FROM N.A., AND PARTIAL SEQUENCE.
 RC TISSUE=Liver;
 RX MEDLINE=95154329; PubMed=7851425;

RA de Bise D., Baira D., Simmaco M., John R.A., Bossa F.;
 RT "Primary structure and tissue distribution of human 4-aminobutyrate
 aminotransferase.";
 RL Eur. J. Biochem. 227:476-480(1995).
 RN [3]
 RP VARIANT LYS-220
 RX MEDLINE=99336116; PubMed=10407778;
 RA Medina-Kaue L.R., Tobin A.J., De Melreir L., Jaeken J., Jakobs C.,
 RT Nyhan W.L., Gibson K.M.;
 RL "4-aminobutyrate aminotransferase (GABA-transaminase) deficiency.";
 RL J. Inher. Metab. Dis. 22:414-427(1999).
 CC -1- CATALYTIC ACTIVITY: 4-aminobutanoate + 2-oxoglutarate = succinate
 CC semialdehyde + L-glutamate.
 CC -1- COFACTOR: PYRIDOXAL PHOSPHATE.
 CC -1- SUBUNIT: MONOMER (PROBABLE).
 CC -1- SUBCELLULAR LOCATION: Mitochondrial matrix.
 CC -1- TISSUE SPECIFICITY: LIVER > PANCREAS > BRAIN > KIDNEY > HEART >
 CC PLACENTA.
 CC -1- DISEASE: DEFECTS IN ABAT ARE A CAUSE OF GABA-AT DEFICIENCY WHOSE
 CC PHENOTYPE INCLUDES PSYCHOMOTOR RETARDATION, HYPOTONIA,
 CC HYPERREFLEXIA, LETHARGY, REFRACTORY SEIZURES, AND EEG
 CC ANOMALITIES.
 CC -1- SIMILARITY: BELONGS TO CLASS-III OF PYRIDOXAL-PHOSPHATE-DEPENDENT
 CC AMINOTRANSFERASES.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: L32961; AAA74449.1; -
 DR HSSP: P80147; 1GTX.
 DR MIM: 137150; -
 DR InterPro: IPR000954; Aminotran_3.
 DR Pfam: PF00202; aminotran_3; 1.
 DR PROSITE: PS00600; AA-TRANSFER_CLASS_3; 1.
 KW Transferrase; Aminotransferase; Pyridoxal phosphate;
 KW Neurotransmitter degradation; Mitochondrion; Transit peptide;
 KW Disease mutation.
 FT TRANSIT 1 28 MITOCHONDRION.
 FT CHAIN 29 500 4-AMINO-BUTYRATE AMINOTRANSFERASE.
 FT BINDING 357 357 PYRIDOXAL PHOSPHATE.
 FT VARIANT 220 220 R->K (IN GABA-AT DEFICIENCY; 25%
 REDUCTION IN ACTIVITY).
 FT FTID=VAR_008883.
 FT D->H (IN REF. 2).
 FT E->L (IN REF. 2).
 FT E->G (IN REF. 2).
 FT K->O (IN REF. 2).
 FT W->G (IN REF. 2).
 FT S->A (IN REF. 2).
 FT G->R (IN REF. 2).
 FT C->G (IN REF. 2).
 FT L->H (IN REF. 2).
 FT SEQUENCE 500 AA; 56557 MW; 41199085693F80AD CRC64;
 Query Match 38.8%; Score 41.5; DB 1; Length 500;
 Best Local Similarity 55.6%; Pred. No. 41;
 Matches 10; Conservative 4; Mismatches 3; Indels 1; Gaps 1;
 QY 3 TFOVEVPSGSHDSQKA 20
 I:::IIII:II I I
 DB 17 TYRLVPSGRHI-SQAAA 33
 RESULT 14
 RFH_ECOLI
 ID RFH_ECOLI STANDARD; PRT; 141 AA.
 AC P28369; P77246;

DT 01-DEC-1992 (Rel. 24, Created)
 DT 01-DEC-1992 (Rel. 24, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Peptide chain release factor homolog (RF-H).
 GN PRFH OR B0236.
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 CC Escherichia.
 NX NCBI_TaxID=562;
 RX [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12;
 RX MEDLINE=90330577; PubMed=1695895;
 RA Henrich B., Monnerjahn U., Plapp R.;
 RT "Peptidase D gene (pepD) of Escherichia coli K-12: nucleotide
 RT sequence, transcript mapping, and comparison with other peptidase
 RT genes.";
 RL J. Bacteriol. 172:4641-4651(1990).
 RN [2]
 RN SEQUENCE FROM N.A.
 RP STRAIN=K12 / MG1655;
 RX MEDLINE=97426617; PubMed=9278503;
 RA Blatner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Kose D.J.,
 RA Mau B., Shao Y.;
 RL "The complete genome sequence of Escherichia coli K-12.";
 RL Science 277:1453-1474(1997).
 RN [3]
 RN SEQUENCE FROM N.A.
 RP STRAIN=K12 / W3110;
 RA Takemoto K., Mori H., Murayama N., Kataoka K., Yano M., Itoh T.,
 RA Yamamoto Y., Inokuchi H., Miki T., Hatada E., Fukuda R., Ichihara S.,
 RA Mizuno T., Makino K., Nakata A., Yura T., Sempel G., Mizobuchi K.;
 RT "Systematic sequencing of the Escherichia coli genome: analysis of the
 RT 4.0 - 6.0 min (189,987 - 281,416bp) region.";
 RL Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RN SEQUENCE FROM N.A.
 RP Schramm S., Duncan M., Allen E., Araujo R., Aparicio A., Chung E.,
 RA Davis K., Federpiet N., Hyman R., Kalman S., Komp C., Kurd O.,
 RA Lashkari D., Lew H., Lin D., Namath A., Oetner P., Roberts D.,
 RA Davis R.W.;
 RL Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RN IDENTIFICATION OF PROTEIN.
 RP MEDLINE=93027135; PubMed=1408743;
 RX Pel H.J., Rep M., Grivell L.A.;
 RT "Sequence comparison of new prokaryotic and mitochondrial members of
 RT the polypeptide chain release factor family predicts a five-domain
 RL model for release factor structure.";
 RL Nucleic Acids Res. 20:4423-4428(1992).
 CC -1- FUNCTION: COULD BE PROTEIN FACTOR INVOLVED IN TRANSLATIONAL
 CC TERMINATION.
 CC -1- SIMILARITY: BELONGS TO THE PROKARYOTIC AND MITOCHONDRIAL RELEASE
 CC FACTORS FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M34034; -; NOT_ANNOTATED_CDS.
 DR EMBL: AE000132; AAC73340.1; -
 DR EMBL: DB3536; BAA77905.1; -
 DR EMBL: U70214; AAB08656.1; -
 DR Ecogene: EG11496; prfh.
 DR InterPro: IPR000352; pep_rel_factor_I.
 DR Pfam: PF00472; RF-I; 1.
 DR PROSITE: PS00745; RF_PROK_I; 1.

KW Protein biosynthesis; Complete proteome.
 FT DOMAIN 104 107 POLY-GLN.
 FT CONFLICT 1 6 MGIRK -> MEETGRSDTLRSALVSLDGDNMAISES
 FT (IN REF. 2 AND 4).
 SQ SEQUENCE 141 AA; 16177 MW; 9D2BD9AF7A04831 CRC64;

Query Match 38.3%; Score 41; DB 1; Length 141;
 Best Local Similarity 35.0%; Pred. No. 13;
 Matches 7; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

OY 2 ETVQYEVPGSOHIDSQKAI 21
 11 : 11 11 : 1 :
 DB 49 ETLRSSGPGQHVHKTDSAV 68

RESULT 15

YP68_CAEEL STANDARD; PRT; 313 AA.
 ID YP68_CAEEL
 AC 009217;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical 37.0 kDa protein B0495.8 in chromosome II.
 GN B0495.8.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RA Kirsten J.;
 RL Submitted (FEB-1995) to the EMBL/GenBank/DBJ databases.
 -I- SIMILARITY: SOME. TO YEAST YDL087C AND S.POMBE SPCC16A11.13.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; U21317; AAA62527.1; -
 DR Wormpep; B0495.8; CE01766.
 KW Hypothetical protein.
 SQ SEQUENCE 313 AA; 36977 MW; 000D2327621BFED0 CRC64;

Query Match 38.3%; Score 41; DB 1; Length 313;
 Best Local Similarity 46.2%; Pred. No. 30;
 Matches 6; Conservative 6; Mismatches 1; Indels 0; Gaps 0;
 OY 7 ETVGSOHIDSOKK 19
 :: 1111:1:::1
 DB 11 QLMGSOHVDNKKR 23

Search completed: July 3, 2002, 09:04:43
 Job time: 583 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 3, 2002, 09:06:05 ; Search time 74.17 Seconds
(without alignments)
48.981 Million cell updates/sec

Title: US-09-786-648-5
Perfect score: 107
Sequence: 1 GETFQVEVPGSHIDSQKKAI 21

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues
Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 08
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL.19:*
1: sp.archaea:*
2: sp.bacteria:*
3: sp.fungi:*
4: sp.human:*
5: sp.invertebrate:*
6: sp.mammal:*
7: sp.mhc:*
8: sp.organelle:*
9: sp.phage:*
10: sp.plant:*
11: sp.rodent:*
12: sp.virus:*
13: sp.vertibrate:*
14: sp.unclassified:*
15: sp.virus:*
16: sp.bacteriophage:*
17: sp.archaeo:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	101	94.4	124	2	057193
2	101	94.4	124	2	09RP15
3	101	94.4	124	2	056635
4	101	94.4	124	2	093V32
5	101	94.4	124	2	094M01
6	95	88.8	103	2	09R646
7	54	50.5	508	4	099987
8	52	48.6	552	10	09LJB0
9	49	45.8	255	10	09C815
10	49	45.8	257	10	09C875
11	49	45.8	428	10	09LGM2
12	47	43.9	352	10	040742
13	47	43.9	395	5	09NKR5
14	46	43.0	91	10	09SVF4
15	46	43.0	574	10	09SXC9
16	45.5	42.5	1166	3	09PAE0

17	45	42.1	439	10	09LY67	091Y67 arabidopsis
18	45	42.1	2732	12	09J3F2	09J3F2 murine hepa
19	45	42.1	2733	12	09PYA2	09PYA2 murine hepa
20	45	42.1	2733	12	09J226	09J226 murine hepa
21	45	42.1	2733	12	09J3E8	09J3E8 murine hepa
22	44.5	41.6	565	10	022511	022511 vitis vinif
23	44	41.1	293	10	049876	049876 lupinus alb
24	44	41.1	330	11	09QZL7	09QZL7 mus musculu
25	44	41.1	374	5	09U3M9	09U3M9 caenorhabdi
26	44	41.1	467	6	09N136	09N136 ovis arie
27	44	41.1	467	11	091VD0	091VD0 mus musculu
28	44	41.1	918	3	09P606	09P606 neurospora
29	43.5	40.7	641	10	09SKB2	09SKB2 arabidopsis
30	43.5	40.7	641	10	093Z40	093Z40 arabidopsis
31	43	40.2	356	5	09NF90	09NF90 leishmania
32	43	40.2	368	10	09SNA4	09SNA4 arabidopsis
33	43	40.2	843	12	09QMN7	09QMN7 hepatitis b
34	43	40.2	843	12	091817	091817 hepatitis b
35	43	40.2	849	16	0981D6	0981D6 rhizobium l
36	43	40.2	1714	13	090995	090995 gallus gall
37	43	40.2	1810	13	090824	090824 gallus gall
38	42	39.3	204	16	09H7A0	09H7A0 pseudomonas
39	42	39.3	330	16	099VM1	099VM1 staphylococ
40	42	39.3	427	13	042099	042099 cyprinus ca
41	42	39.3	672	16	09A818	09A818 caulobacter
42	42	39.3	697	5	0965W6	0965W6 caenorhabdi
43	42	39.3	802	10	0947W6	0947W6 oryza sativ
44	42	39.3	945	10	094DZ7	094DZ7 oryza sativ
45	42	39.3	946	10	094E00	094E00 oryza sativ

ALIGNMENTS

RESULT 1
057193 PRELIMINARY; PRT; 124 AA.
AC 057193;
DT 01-NOV-1996 (TREMREL. 01, Created)
DT 01-NOV-1996 (TREMREL. 01, Last sequence update)
DT 01-JUN-2001 (TREMREL. 17, Last annotation update)
DE CHOLERA TOXIN B PROTEIN (CTB) PRECURSOR (CTB).
GN CTXB.
OS Vibrio cholerae.
OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
OX NCBI_TaxID=666;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CLASSICAL STRAIN 569B;
RX MEDLINE=9135224; PubMed=1883840;
RA Dams E., De Wolf M., Dierick W.;
RT "Nucleotide sequence analysis of the CT operon of the Vibrio cholerae
classical strain 569B.";
RT Biochim. Biophys. Acta 1090:139-141(1991).
RL [2]
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CLASSICAL BIOTYPE 569B;
RA Shi C., Cao C., Zhang J., Ma O.;
RN Chin. Biochem. J. 9:395-399(1993).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=CLASSICAL BIOTYPE 569B;
RA Xu L.;
RT Submitted (APR-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL: X58785; CAA41591.1; -;
DR EMBL: U25679; AAC34728.1; -;
DR EMBL: A00931; CAA00098.1; -;
DR HSSP: P01556; ZCHB.
DR InterPro: IPR001835; Enterotoxin_B.
DR Pfam: PF01376; Enterotoxin_B.1.
DR PRINTS: PR00772; ENTEROTOXINB.
DR ProDom: PD012805; Enterotoxin_B.1.
KW Signal.

DR HSSP; P01556; 2CHB.

RN [1]

DE PEPTIDE CHAIN RELEASE FACTOR, POTATIVE
GN F10C21.2.
OS Arabidopsis thaliana (Mouse-ear cress)

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC Euphorbia; Brassicales; Brassicaceae; Arabidopsis.
 NCBI_TaxID=3702;
 [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. COLUMBIA;
 RX MEDLINE=21016719; PubMed=11130712;
 RA Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,
 RA White O., Alonso J., Altati H., Araujo R., Bowman C.L., Brooks S.Y.,
 RA Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,
 RA Chung M.K., Conn L., Conway A.B., Conway A.R., Cressy T.H., Dewar K.,
 RA Dunn P., Egu P., Feldhym T.V., Feng J.-D., Feng B., Fujii C.Y.,
 RA Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huizer L.,
 RA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,
 RA Kim C.J., Koo H.L., Kremenetska I., Kurtz D.B., Kwan A., Lam B.,
 RA Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,
 RA Lin X., Liu S.X., Liu Z.A., Lueros J.S., Maiti R., Marziani A.,
 RA Miltischer J., Miranda M., Nguyen M., Niernan W.C., Osborne B.I.,
 RA Pai G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,
 RA Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M.,
 RA Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,
 RA Utterback T., Van Aken S., Vaysberg M., Vysotskaia V.S., Walker M.,
 RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.,
 RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis
 thaliana."
 RL Nature 408:816-820(2000).
 DR EMBL; AC051630; AAG51209.1; -
 DR InterPro; IPR000352; Pep_rel_factor_1.
 DR Pfam; PF00472; RF-1; 1
 SO SEQUENCE 255 AA; 27445 MW; B5AF38991FF0D0F8 CRC64;

Query Match 45.8%; Score 49; DB 10; Length 255;
 Best Local Similarity 45.0%; Pred. No. 3.9;
 Matches 9; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

OY 2 ETRVPSGQHIDSOKKAI 21
 DB 101 ETRVPSGQGHNRKDSAV 120

RESULT 10
 OY 09C875 PRELIMINARY; PRT; 257 AA.
 AC 09C875;
 DT 01-JUN-2001 (TREMBLrel. 17, Created)
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE HYPOTHETICAL 27.6 KDA PROTEIN.
 GN T1609.18.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC Euphorbia; Brassicales; Brassicaceae; Arabidopsis.
 NCBI_TaxID=3702;
 [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. COLUMBIA;
 RX MEDLINE=21016719; PubMed=11130712;
 RA Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,
 RA White O., Alonso J., Altati H., Araujo R., Bowman C.L., Brooks S.Y.,
 RA Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,
 RA Chung M.K., Conn L., Conway A.B., Conway A.R., Cressy T.H., Dewar K.,
 RA Dunn P., Egu P., Feldhym T.V., Feng J.-D., Feng B., Fujii C.Y.,
 RA Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huizer L.,
 RA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,
 RA Kim C.J., Koo H.L., Kremenetska I., Kurtz D.B., Kwan A., Lam B.,
 RA Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,
 RA Lin X., Liu S.X., Liu Z.A., Lueros J.S., Maiti R., Marziani A.,
 RA Miltischer J., Miranda M., Nguyen M., Niernan W.C., Osborne B.I.,
 RA Pai G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,
 RA Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M.,

RA Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,
 RA Utterback T., Van Aken S., Vaysberg M., Vysotskaia V.S., Walker M.,
 RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.,
 RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis
 thaliana."
 RL Nature 408:816-820(2000).
 DR EMBL; AC027035; AAG51290.1; -
 DR InterPro; IPR000352; Pep_rel_factor_1.
 DR Pfam; PF00472; RF-1; 1
 KW Hypothetical protein.
 SO SEQUENCE 257 AA; 27645 MW; 9F85B09C029E97C9 CRC64;

Query Match 45.8%; Score 49; DB 10; Length 257;
 Best Local Similarity 45.0%; Pred. No. 3.9;
 Matches 9; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

OY 2 ETRVPSGQHIDSOKKAI 21
 DB 101 ETRVPSGQGHNRKDSAV 120

RESULT 11
 OY 09LGM2 PRELIMINARY; PRT; 428 AA.
 AC 09LGM2;
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE ESTS A0056822(S20908).
 OS Oryza sativa (Rice).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoideae; Oryzaceae; Oryza.
 NCBI_TaxID=4530;
 [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. NIPPONBARE;
 RA Sasaki T., Matsumoto T., Yamamoto K.;
 RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC
 clone: P0041E11."
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. NIPPONBARE;
 RA Sasaki T., Matsumoto T., Yamamoto K.;
 RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC
 clone: P0433P09."
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AP002521; BAA96762.1; -
 DR EMBL; AP002539; BAA08201.1; -
 DR HSSP; 006319; 1BUC.
 DR InterPro; IPR001552; Acyl-CoA_dh.
 DR Pfam; PF00441; Acyl-CoA_dh.1.
 DR Pfam; PF02770; Acyl-CoA_dh.M.1.
 DR Pfam; PF02771; Acyl-CoA_dh.N.1.
 DR PROSITE; PS00073; ACYL_COA_DH.2; UNKNOWN.1.
 SO SEQUENCE 428 AA; 46132 MW; 8D34E369A88E6367 CRC64;

Query Match 45.8%; Score 49; DB 10; Length 428;
 Best Local Similarity 52.6%; Pred. No. 7;
 Matches 10; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

OY 3 ETRVPSGQHIDSOKKAI 21
 DB 183 TATKVPGGWHDGQKRWI 201

RESULT 12
 ID 040742 PRELIMINARY; PRT; 392 AA.
 AC 040742;

Derivatives of *Escherichia coli* heat labile enterotoxins useful as

XX Claim 1; Page 13; 62pp; English.
PS
XX
CC The invention relates to peptide fragments of the Escherichia coli heat
CC labile enterotoxin (Etx) and its closely related homologue, cholera
CC toxin (Ctx) from Vibrio cholerae which do not bind to the ubiquitous
CC GM-1 ganglioside receptors on host cell surfaces. Both Etx and Ctx are
CC composed of one A subunit and five identical B subunits. The A subunit
CC is responsible for toxicity, possessing adenosine diphosphate (ADP) ADP-
CC ribosyltransferase activity, while the B subunits (EtxB and CtxB)
CC facilitate the entry of subunit A into the host cell via the binding and
CC cross-linking of GM-1 receptors. Although GM-1 binding is responsible
CC for some of the effects of Etx and Ctx, it has been found that certain
CC effects of the toxins, such as immunomodulation, are not mediated
CC through GM-1 binding. The peptides of the invention are fragments of the
CC beta-4- alpha-2 loop of EtxB and/or CtxB, exerting the same effects as
CC normal EtxB and CtxB subunits, except that they do not bind or cross link
CC GM-1. They may be used in medicine as an immunomodulator or adjuvant.
CC They may also be used as an inhibitor for toxin-induced diarrhoea.
CC Therefore, the peptides may be used in the production of a composition
CC for treating, preventing and/or modulating a disease associated with an
CC immune disorder and/or toxin-induced diarrhoea. Sequences AAY87460-Y87463
CC represent preferred peptides of the invention, AAY87460 being
CC particularly preferred.
XX
XX
SQ Sequence 7 AA;

Query Match 100.0%; Score 39; DB 21; Length 7;
Best Local Similarity 100.0%; Pred. No. 6.4e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EVPGSQH 7
| | | | | | |
Db 1 evpsqgh 7

RESULT 2
AAY87464
ID AAY87464 standard; peptide; 8 AA.
XX

AC AAY87464;

DT 03-JUL-2000 (first entry)

DE Cholera toxin B/enterotoxin B-derived immunomodulatory peptide.

KW Cholera toxin subunit B; CtxB; heat labile enterotoxin subunit B; EtxB;

KM beta-4-alpha-2 loop; GM-1 ganglioside receptor; immunomodulation;

KW adjuvant; immune disorder; diarrhoea.

OS Vibrio cholerae.

OS Escherichia coli.

PN WO200014114-A1.

PD 16-MAR-2000.

PF 07-SEP-1999; 99WO-GB02970.

PR 07-SEP-1998; 98GB-0019484.

PA (UYBR-) UNIV BRISTOL.

PI Williams NA, Hirst TR;

PT WPI; 2000-256943/22.

XX Derivatives of Escherichia coli heat labile enterotoxins useful as
XX immunomodulators and for treating diarrhoea and which do not bind the
XX glycolipid receptor GM-1 -
PS Example 5; Page 45; 62pp; English.

XX
CC The invention relates to peptide fragments of the Escherichia coli heat
CC labile enterotoxin (Etx) and its closely related homologue, cholera
CC toxin (Ctx) from Vibrio cholerae which do not bind to the ubiquitous
CC GM-1 ganglioside receptors on host cell surfaces. Both Etx and Ctx are
CC composed of one A subunit and five identical B subunits. The A subunit
CC is responsible for toxicity, possessing adenosine diphosphate (ADP) ADP-
CC ribosyltransferase activity, while the B subunits (EtxB and CtxB)
CC facilitate the entry of subunit A into the host cell via the binding and
CC cross-linking of GM-1 receptors. Although GM-1 binding is responsible
CC for some of the effects of Etx and Ctx, it has been found that certain
CC effects of the toxins, such as immunomodulation, are not mediated
CC through GM-1 binding. The peptides of the invention (AAY87461-Y87463) are
CC fragments of the beta-4-alpha-2 loop of EtxB and/or CtxB, exerting the
CC same effects as normal EtxB and CtxB subunits, except that they do not
CC bind or cross link GM-1. They may be used in medicine as an
CC immunomodulator or adjuvant. They may also be used as an inhibitor for
CC toxin-induced diarrhoea. Therefore, the peptides may be used in the
CC production of a composition for treating, preventing and/or modulating a
CC disease associated with an immune disorder and/or toxin-induced
CC diarrhoea. Sequences AAY87464-Y87465 represent peptides used in an
CC exemplification of the present invention to assess whether a peptide
CC corresponding to a portion of the beta-4-alpha-2 loop of EtxB and CtxB
CC has immunomodulatory effects. Peptide AAY87464 corresponds to residues
CC 51-58 of the EtxB/CtxB beta-4-alpha-2 loop, and peptide AAY87465 is a
CC randomly selected control peptide.
XX
XX
SQ Sequence 8 AA;

Query Match 100.0%; Score 39; DB 21; Length 8;
Best Local Similarity 100.0%; Pred. No. 6.4e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EVPGSQH 7
| | | | | | |
Db 1 evpsqgh 7

RESULT 3
AAY87461
ID AAY87461 standard; peptide; 12 AA.
XX

AC AAY87461;

DT 03-JUL-2000 (first entry)

DE Cholera toxin B/enterotoxin B-derived peptide, SEQ ID NO:3.

KW Cholera toxin subunit B; CtxB; heat labile enterotoxin subunit B; EtxB;

KM beta-4-alpha-2 loop; GM-1 ganglioside receptor; immunomodulation;

KW adjuvant; immune disorder; diarrhoea.

OS Vibrio cholerae.

OS Escherichia coli.

PN WO200014114-A1.

PD 16-MAR-2000.

PF 07-SEP-1999; 99WO-GB02970.

PR 07-SEP-1998; 98GB-0019484.

PA (UYBR-) UNIV BRISTOL.

PI Williams NA, Hirst TR;

PT WPI; 2000-256943/22.

XX Derivatives of Escherichia coli heat labile enterotoxins useful as
XX immunomodulators and for treating diarrhoea and which do not bind the
XX glycolipid receptor GM-1 -
PT

XX Disclosure: Page 15; 62pp; English.
PS
XX
CC The invention relates to peptide fragments of the Escherichia coli heat
CC labile enterotoxin (Etx) and its closely related homologue, cholera
CC toxin (Ctx) from Vibrio cholerae which do not bind to the ubiquitous
CC GM-1 ganglioside receptors on host cell surfaces. Both Etx and Ctx are
CC composed of one A subunit and five identical B subunits. The A subunit
CC is responsible for toxicity, possessing adenosine diphosphate (ADP)-
CC ribosyltransferase activity, while the B subunits (EtxB and CtxB)
CC facilitate the entry of subunit A into the host cell via the binding and
CC cross-linking of GM-1 receptors. Although GM-1 binding is responsible
CC for some of the effects of Etx and Ctx, it has been found that certain
CC effects of the toxins, such as immunomodulation, are not mediated
CC through GM-1 binding. The peptides of the invention are fragments of the
CC beta-4- alpha-2 loop of EtxB and/or CtxB, exerting the same effects as
CC normal EtxB and CtxB subunits, except that they do not bind or cross link
CC GM-1. They may be used in medicine as an immunomodulator or adjuvant.
CC They may also be used as an inhibitor for toxin-induced diarrhoea.
CC Therefore, the peptides may be used in the production of a composition
CC for treating, preventing and/or modulating a disease associated with an
CC immune disorder and/or toxin-induced diarrhoea. Sequences AAY87460-Y87463
CC represent preferred peptides of the invention, AAY87460 being
CC particularly preferred.
SQ Sequence 12 AA:

Query Match 100.0%; Score 39; DB 21; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.29; 0; Indels 0; Gaps 0;
Matches 7; Conservative 0; Mismatches 0;

OY 1 EYPGSGH 7
|||||||
DB 2 evpgsqh 8

RESULT 4
AAP93498
ID AAP93498 standard; protein; 15 AA.
XX
AC AAP93498;
XX
DT 03-MAY-1990 (first entry)
XX
DE CTP3 epitope of the Cholera toxin B subunit.
XX
KW CTP3 epitope of Cholera toxin B subunit; flagellin fusion protein;
KM vaccine; immunotherapy; ds;
XX
PN WO8910967-A.
XX
PD 16-NOV-1989.
XX
PF 05-MAY-1989; 89WO-US01932.
XX
PR 05-MAY-1988; 88US-0190570.
XX
PA (PRAX-) PRAXIS BIOLOGICS INC.
PA (STRD) LEYLAND STANFORD JUNIOR UNIV.
XX
PI Marjarian WR, Stocker BAD, Newton SMC;
XX
DR WPI, 1989-356496/48.
DR N-PSDB: AAN92414.
XX
PT New recombinant flagellin gene including sequence - for heterologous
PT epitope, and expressed fusion proteins, useful in vaccines and for prodn.
XX of antibodies;
XX
PS Disclosure; fig.4B; 137pp; English.
XX This sequence corresponds to the CTP3 epitope of the Cholera toxin B

CC subunit. The DNA sequence encoding this ligates to othersynthetic
CC oligonucleotides to form a new recombinant gene. This encodes
CC a flagellin fusion protein which can be used in vaccines for immuno-
CC therapy.
XX
SQ Sequence 15 AA:

Query Match 100.0%; Score 39; DB 10; Length 15;
Best Local Similarity 100.0%; Pred. No. 0.37;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EYPGSGH 7
|||||||
DB 2 evpgsqh 8

RESULT 5
AAR85125
ID AAR85125 standard; peptide; 15 AA.
XX
AC AAR85125;
XX
DT 13-JUN-1996 (first entry)
XX
DE Cholera toxin B antigenic peptide fragment CTP3.
XX
KW Conjugate; cholera; B toxin; peptide fragment; microparticulate;
KW inert carrier; modified silica; thyroglobulin; oral vaccine;
KW immunisation; infection; insoluble; digestive tract; antigen;
KW intestines; antibodies; secretory; IGA class.
XX
OS Vibrio cholerae.
XX
PN WO9529701-A1.
XX
PD 09-NOV-1995.
XX
PF 02-MAY-1995; 95WO-EP01661.
XX
PR 03-MAY-1994; 94IL-0109519.
XX
PA (YEDA) YEDA RES & DEV CO LTD.
XX
PI Marks RS, Mirelman D, Sela M;
XX
DR WPI, 1995-403805/51.
XX
DE Vaccines for oral immunisation against infecting agents, e.g.
PT cholera - comprise a conjugate of an antigen of an infecting agent
PT covalently bound to micro:particulate inert carrier, e.g. modified
PT aldehyde silica
XX
PS Claim 7; Page 25; 40pp; English.
XX
CC A compsn. comprising a conjugate of an antigenic cholera B toxin
CC peptide fragment (e.g. AAR85126-30 and esp. AAR85125), covalently
CC bound to a microparticulate inert carrier (e.g. modified silica or
CC thyroglobulin) can be used as an oral vaccine for immunisation
CC against cholera infection. The inert carrier is insoluble in the
CC digestive tract, allowing presentation of the antigen in the
CC intestines, where it will elicit antibodies mainly of the
CC secretory IGA class.
XX
SQ Sequence 15 AA:

Query Match 100.0%; Score 39; DB 16; Length 15;
Best Local Similarity 100.0%; Pred. No. 0.37;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EYPGSGH 7
|||||||

Db 2 evpsqgh 8

RESULT 6
AA87462
ID AAY87462 standard; peptide: 21 AA.
XX
XX AAY87462;
XX
XX 03-JUL-2000 (first entry)
XX
XX Cholera toxin B/enterotoxin B-derived peptide, SEQ ID NO:4.
XX
XX Cholera toxin subunit B; CtxB; heat labile enterotoxin subunit B; EtxB;
XX beta-4-alpha-2 loop; GM-1 ganglioside receptor; immunomodulation;
XX adjuvant; immune disorder; diarrhoea.
XX
XX Vibrio cholerae;
XX Escherichia coli.
XX
XX WO200014114-A1.
XX
XX 16-MAR-2000.
XX
XX 07-SEP-1999; 99WO-G802970.
XX
XX 07-SEP-1998; 98GB-0019484.
XX
XX (UYBR-) UNIV BRISTOL.
XX
XX Williams NA, Hirst TR;
XX
XX WPI: 2000-256943/22.
XX
XX Derivatives of Escherichia coli heat labile enterotoxins useful as
PT immunomodulators and for treating diarrhea and which do not bind the
PT glycolipid receptor GM-1 -
XX
XX
XX Disclosure: Page 15; 62pp; English.

XX The invention relates to peptide fragments of the Escherichia coli heat
CC labile enterotoxin (Etx) and its closely related homologue, cholera
CC toxin (Ctx) from Vibrio cholerae which do not bind to the ubiquitous
CC GM-1 ganglioside receptors on host cell surfaces. Both Etx and Ctx are
CC composed of one A subunit and five identical B subunits. The A subunit
CC is responsible for toxicity, possessing adenosine diphosphate (ADP) ADP-
CC ribosyltransferase activity, while the B subunits (EtxB and CtxB)
CC facilitate the entry of subunit A into the host cell via the binding and
CC cross-linking of GM-1 receptors. Although GM-1 binding is responsible
CC for some of the effects of Etx and Ctx, it has been found that certain
CC effects of the toxins, such as immunomodulation, are not mediated
CC through GM-1 binding. The peptides of the invention are fragments of the
CC beta-4- alpha-2 loop of EtxB and/or CtxB, exerting the same effects as
CC normal EtxB and CtxB subunits, except that they do not bind or cross link
CC GM-1. They may be used in medicine as an immunomodulator or adjuvant.
CC They may also be used as an inhibitor for toxin-induced diarrhoea.
CC Therefore, the peptides may be used in the production of a composition
CC for treating, preventing and/or modulating a disease associated with an
CC immune disorder and/or toxin-induced diarrhoea. Sequences AAY87460-Y87463
CC represent preferred peptides of the invention, AAY87460 being
CC particularly preferred.
XX
XX
XX Sequence 21 AA;
SQ

Query Match 100.0%; Score 39; DB 21; Length 21;
Best Local Similarity 100.0%; Pred. No. 0.52;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EVPSQGH 7
| | | | | | |
Db 7 evpsqgh 13

RESULT 7
AA87463
ID AAY87463 standard; peptide: 21 AA.
XX
XX AAY87463;
XX
XX 03-JUL-2000 (first entry)
XX
XX E. coli heat labile enterotoxin B-derived peptide, SEQ ID NO:5.
XX
XX Heat labile enterotoxin subunit B; EtxB;
XX beta-4-alpha-2 loop; GM-1 ganglioside receptor; immunomodulation;
XX adjuvant; immune disorder; diarrhoea.
XX
XX Escherichia coli.
XX
XX WO200014114-A1.
XX
XX 16-MAR-2000.
XX
XX 07-SEP-1999; 99WO-G802970.
XX
XX 07-SEP-1998; 98GB-0019484.
XX
XX (UYBR-) UNIV BRISTOL.
XX
XX Williams NA, Hirst TR;
XX
XX WPI: 2000-256943/22.
XX
XX Derivatives of Escherichia coli heat labile enterotoxins useful as
PT immunomodulators and for treating diarrhea and which do not bind the
PT glycolipid receptor GM-1 -
XX
XX
XX Disclosure: Page 15; 62pp; English.

XX The invention relates to peptide fragments of the Escherichia coli heat
CC labile enterotoxin (Etx) and its closely related homologue, cholera
CC toxin (Ctx) from Vibrio cholerae which do not bind to the ubiquitous
CC GM-1 ganglioside receptors on host cell surfaces. Both Etx and Ctx are
CC composed of one A subunit and five identical B subunits. The A subunit
CC is responsible for toxicity, possessing adenosine diphosphate (ADP) ADP-
CC ribosyltransferase activity, while the B subunits (EtxB and CtxB)
CC facilitate the entry of subunit A into the host cell via the binding and
CC cross-linking of GM-1 receptors. Although GM-1 binding is responsible
CC for some of the effects of Etx and Ctx, it has been found that certain
CC effects of the toxins, such as immunomodulation, are not mediated
CC through GM-1 binding. The peptides of the invention are fragments of the
CC beta-4- alpha-2 loop of EtxB and/or CtxB, exerting the same effects as
CC normal EtxB and CtxB subunits, except that they do not bind or cross link
CC GM-1. They may be used in medicine as an immunomodulator or adjuvant.
CC They may also be used as an inhibitor for toxin-induced diarrhoea.
CC Therefore, the peptides may be used in the production of a composition
CC for treating, preventing and/or modulating a disease associated with an
CC immune disorder and/or toxin-induced diarrhoea. Sequences AAY87460-Y87463
CC represent preferred peptides of the invention, AAY87460 being
CC particularly preferred.
XX
XX
XX Sequence 21 AA;
SQ

Query Match 100.0%; Score 39; DB 21; Length 21;
Best Local Similarity 100.0%; Pred. No. 0.52;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EVPSQGH 7
| | | | | | |
Db 7 evpsqgh 13

RESULT 8
AAR76748

ID AAR76748 standard; Protein; 23 AA.
 XX AAR76748;
 AC
 DT 18-MAR-1996 (first entry)
 XX
 DE Residues 50-64 of cholera toxin B subunit and FimH 224-226.
 XX
 KW FimH; type 1 fimbriae; organelle; adhesin; alpha-D-mannoside residue;
 KW FimH; FimH; receptor binding site; PCR; amplify; ss.
 XX
 OS Chimeric - Vibrio cholerae.
 OS Chimeric - Escherichia coli.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 1 /note= "Represents FimH residue 224"
 FT Peptide 2..4 /note= "Linker peptide"
 FT Peptide 5..19 /note= "Cholera toxin B subunit 50-64"
 FT Peptide 20..22 /note= "Linker peptide"
 FT Misc-difference 23 /note= "Represents FimH residue 226"
 FT
 XX
 PN WO9520657-A1.
 PD 03-AUG-1995.
 XX
 PF 27-JAN-1995; 95WO-DK00042.
 XX
 PR 27-JAN-1994; 94US-0187166.
 XX
 PA (GXBI-) GX BIOSYSTEMS AS.
 XX
 PI Hasty DL, Klemm P, Molin S, Pallesen L, Sokurenko EV;
 XX
 DR WPI; 1995-275442/36.
 DR N-PSDB; AAQ93061.
 XX
 PT Receptor specific bacterial adhesins - useful for targetting active
 PT compounds and microbial cells to locations of receptors
 PS
 XX Example 3; Page 58; 152pp; English.
 XX
 CC This sequence is encoded by a fragment of the the plasmid pLP93
 CC which was used in the production of fimb fusion genes comprising
 CC the cholera toxin B subunit inserted into the fimb gene. This insert
 CC shows the inclusion of the B subunit into the fimb protein at position
 CC 224-226. The chimeric genes were then opt. further modified by insertion
 CC of the hepatitis B virus surface antigen pre-S2 region into a different
 CC position of the fimb adhesin of type 1 fimbriae. Restriction site handles
 CC (BglII-sites) were introduced into the fimb gene, and the foreign
 CC epitopes are then inserted in-frame. In the selected positions the
 CC insertion of the epitopes did not significantly alter the adhesive
 CC function of the fimb protein. The expression of the chimeric proteins
 CC on the surface of fimbriae on bacterial hosts illustrated the possibility
 CC of using bacterial adhesins as general presenters of foreign antigens and
 CC epitopes. These chimeric genes may be used in the production of variant
 CC fimb adhesins which may be useful for targetting active compounds
 CC and microbial cells to locations comprising selected receptors to which
 CC the adhesins bind.
 XX
 SQ Sequence 23 AA;

Query Match 100.0%; Score 39; DB 16; Length 23;
 Best Local Similarity 100.0%; Pred. No. 0.57;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 EVPGSOH 7
 |||||||

DB 6 evpgsqh 12
 RESULT 9
 ID AAP30265 standard; Protein; 26 AA.
 XX AAP30265;
 AC
 DT 21-APR-1992 (first entry)
 XX
 DE Sequence of amino acids 50-75 of the cholera toxin B1 subunit which
 DE carries an Arg at posns. 67 and 73.
 XX
 KW Cholera vaccine; therapy; E.coli infection; enterotoxin LT.
 XX
 OS Vibrio cholerae.
 XX
 PN EP95426-A.
 PD 30-NOV-1983.
 XX
 PF 26-MAY-1983; 83EP-0401052.
 XX
 PR 26-MAY-1982; 82FR-0009167.
 XX
 PA (CNRS) CNRS CENT NAT RECH SCI.
 PA (INSP) INST PASTEUR.
 XX
 PI Milhaud G, Raulais D, Rivallie P, Siffert O, Dodin A;
 XX
 DR WPI; 1983-834645/49.
 XX
 PT Cholera toxin B, sub-unit polypeptide(s) as vaccines and
 PT medicaments - effective against Escherichia coli and Vibrio
 PT cholerae infections, are prepd. by solid phase peptide synthesis
 PS
 XX Claim 7; Page 11; 13pp; French.
 XX
 CC The inventors claim cholera toxin B1 subunit sequences which carry
 CC Arg in posn. 35, 67 or 73 (AAP30264, AAP30265, AAP30600). Tests have
 CC shown that Arg in posns. 35, 67 and 73 play an important role in fixing
 CC these toxins to cell walls. The peptides are used in the treatment of,
 CC and vaccination against, cholera infections and animal and human
 CC infections due to E. coli (enterotoxin LT). The medicament may be
 CC administered by oral, intraperitoneal, sub-cutaneous or intravenous
 CC routes. For vaccines, pref. peptides having 15-30 AAs are injected
 CC without carriers. Unit dose when used as a medicament is 50-500mg as
 CC a vaccine 1-10mg of active cpd.
 XX
 SQ Sequence 26 AA;

Query Match 100.0%; Score 39; DB 4; Length 26;
 Best Local Similarity 100.0%; Pred. No. 0.65;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 EVPGSOH 7
 |||||||
 DB 2 evpgsqh 8

RESULT 10
 ID AAP50439 standard; protein; 41 AA.
 XX AAP50439;
 AC
 DT 01-JAN-1980 (first entry)
 XX
 DE Network polymer which comprises a series of composite E. coli heat-
 DE labile toxin (LT)/heat-stable toxin (ST) polypeptide repeating units.

XX Heat-labile enterotoxin; heat-stable toxin; vaccine; ss.
 XX Synthetic.
 OS
 XX W08502611-A.
 PN
 XX 20-JUN-1985.
 PD
 XX 12-DEC-1984; 84WO-US02030.
 PF
 XX 12-DEC-1983; 83US-0559469.
 PR
 XX (SCRI-) SCRIPPS CLINIC RES.
 PA
 XX Houghten RA;
 PI
 XX WPI: 1985-159230/26.
 DR
 XX
 PT New synthetic polypeptide(s) for part of E. coli enterotoxin(s) -
 useful for vaccination of mammals against the enterotoxin(s)
 PS
 XX Claim 8; Page 100; 120pp; English.
 PS
 XX The repeating units are bonded together by intramolecular
 CC interpolyptide cystine bonds formed between oxidized Cys residues
 CC of the repeating units. This polypeptide may be used in the
 CC vaccination of mammals for protection against the enterotoxins. The
 CC composite polypeptide is made by solid phase synthesis or
 CC recombinant DNA techniques. X9 is Met or Lys, X12 is Ala or Glu,
 CC X39 is Asn or Tyr and X46 is Tyr or Asn. See also AAP50439-57.
 CC
 XX
 SQ Sequence 41 AA;
 OY 1 EVPGSQH 7
 Db 15 evpgsqh 21
 Query Match 100.0%; Score 39; DB 6; Length 41;
 Best Local Similarity 100.0%; Pred. NO. 1;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 11
 AAP50436
 ID AAP50436 standard; protein; 46 AA.
 AC
 XX AAP50436;
 AC
 XX 01-JAN-1980 (first entry)
 DT
 XX
 DE Network polymer which comprises a series of composite E. coli heat-
 DE labile toxin (LT)/heat-stable toxin (ST) polypeptide repeating units.
 DE
 XX Heat-labile enterotoxin; heat-stable toxin; vaccine; ss.
 KW
 XX Synthetic.
 OS
 XX W08502611-A.
 PN
 XX 20-JUN-1985.
 PD
 XX 12-DEC-1984; 84WO-US02030.
 PF
 XX 12-DEC-1983; 83US-0559469.
 PR
 XX (SCRI-) SCRIPPS CLINIC RES.
 PA
 XX Houghten RA;
 PI
 XX WPI: 1985-159230/26.
 DR
 XX

PT New synthetic polypeptide(s) for part of E. coli enterotoxin(s) -
 useful for vaccination of mammals against the enterotoxin(s)
 XX
 PS
 XX Claim 8; Page 100; 120pp; English.
 PS
 XX The repeating units are bonded together by intramolecular
 CC interpolyptide cystine bonds formed between oxidized Cys residues
 CC of the repeating units. This polypeptide may be used in the
 CC vaccination of mammals for protection against the enterotoxins. The
 CC composite polypeptide is made by solid phase synthesis or
 CC recombinant DNA techniques. X9 is Met or Lys, X12 is Ala or Glu,
 CC X39 is Asn or Tyr and X46 is Tyr or Asn. See also AAP50439-57.
 CC
 XX
 SQ Sequence 46 AA;
 OY 1 EVPGSQH 7
 Db 17 evpgsqh 23
 Query Match 100.0%; Score 39; DB 6; Length 46;
 Best Local Similarity 100.0%; Pred. NO. 1.2;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 12
 AAP30600
 ID AAP30600 standard; protein; 47 AA.
 AC
 XX AAP30600;
 AC
 XX 21-APR-1992 (first entry)
 DT
 XX
 DE Sequence of amino acids 350-75 of the cholera toxin B1 subunit which
 DE carries an Arg at posns. 35, 67 and 73.
 DE
 XX Cholera vaccine; therapy; E.coli infection; enterotoxin LT.
 KW
 XX Vibrio cholerae.
 OS
 XX EP95426-A.
 PN
 XX 30-NOV-1983.
 PD
 XX 26-MAY-1983; 83EP-0401052.
 PF
 XX 26-MAY-1982; 82FR-0009167.
 PR
 XX (CNRS) CNRS CENT NAT RECH SCT.
 PA (INSP) INST PASTEUR.
 XX
 PI Milhaud G, Raulais D, Rivaille P, Siffert O, Dodin A;
 PI Guyongruaz A, Delmas A;
 DR WPI: 1983-834645/49.
 DR
 XX Cholera toxin B, sub-unit polypeptide(s) as vaccines and
 PT medicaments - effective against Escherichia coli and Vibrio
 PT cholerae infections, are prepd. by solid phase peptide synthesis
 PT
 XX Claim 8; Page 11; 13pp; French.
 PS
 XX The inventors claim cholera toxin B1 subunit sequences which carry
 CC Arg in posn. 35, 67 or 73 (AAP30264, AAP30265, AAP30600). Tests have
 CC shown that Arg in posns. 35, 67 and 73 play an important role in fixing
 CC these toxins to cell walls. The peptides are used in the treatment of,
 CC and vaccination against, cholera infections and animal and human
 CC infections due to E. coli (enterotoxin LT). The medicament may be
 CC administered by oral, intraperitoneal, sub-cutaneous or intravenous
 CC routes. For vaccines, pref. peptides having 15-30 AAs are injected
 CC without carriers. Unit dose when used as a medicament is 50-500mg as
 CC a vaccine 1-10mg of active cpd.
 CC

SQ Sequence 47 AA;

Query Match 100.0%; Score 39; DB 4; Length 47;
 Best Local Similarity 100.0%; Pred. No. 1.2;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EVPGSOH 7

DB 23 evpgsqh 29

RESULT 13

ID AAR72545 standard; peptide; 93 AA.

AC AAR72545;

DT 28-NOV-1995 (first entry)

DE ADP-ribosylating toxin (verotoxin-1 B-subunit).

KM ADP-ribosylating toxin; pertussis holotoxin; B-subunit;

KM active site; E. coli heat labile toxin; verotoxin-1;

KM Bordetella pertussis vaccines.

OS Bacteria sp.

PN EP646599-A.

PD 05-APR-1995.

PF 23-AUG-1994; 94EP-0306219.

PR 24-AUG-1993; 93US-0110947.

PR 31-MAY-1994; 94US-0251121.

PA (CONN-) CONNAUGHT LAB LTD.

PA (UYAL-) UNIV ALBERTA.

PI Armstrong GD, Cockle SA, Hazes B, Klein MH, Loosmore S;

PI Oomen R, Read RJ, Stein PE;

PS WPI; 1995-132623/18.

CC New modified forms of pertussis holotoxin - developed using

CC crystalline forms of pertussis holotoxin and its complexes with

CC other molecules

CC Disclosure; Fig 5; 54pp; English.

CC AAR72540-R72545 are structurally equivalent B-subunits from three

CC ADP-ribosylating toxins, pertussis holotoxin (PT), E. coli heat

CC labile toxin (LT), and verotoxin-1 (VT). The structural

CC information obtd. from these comparisons was used to identify

CC sites which contribute to PT's biological activity. By modifying

CC these sites the claimed PT mutants of the invention were produced,

CC they can be used in the development of vaccines against Bordetella

CC pertussis infection.

SQ Sequence 93 AA;

Query Match 100.0%; Score 39; DB 16; Length 93;

Best Local Similarity 100.0%; Pred. No. 2.4;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EVPGSOH 7

DB 41 evpgsqh 47

RESULT 14

AAV41816

ID AAV41816 standard; peptide; 93 AA.

AC AAV41816;

DT 08-DEC-1999 (first entry)

DE Escherichia coli verotoxin-1 B-subunit.

KM ADP-ribosylating toxin; PT; crystalline pertussis holotoxin; ETA; DT;

KM three-dimensional structure; LT; immunoprotective; infection.

OS Escherichia coli.

PN US5965385-A.

PD 12-OCT-1999.

PF 06-JUN-1995; 95US-0467974.

PR 22-AUG-1994; 94US-0292968.

PR 24-AUG-1993; 93US-0110947.

PR 31-MAY-1994; 94US-0251121.

PA (CONN-) CONNAUGHT LAB LTD.

PA (UYAL-) UNIV ALBERTA.

PI Read RJ, Stein PE, Cockle SA, Loosmore S, Klein MH, Armstrong GD;

PI Hazes B, Oomen RP;

PS WPI; 1999-579908/49.

CC New method for producing modified pertussis holotoxin -

CC Example 3; Fig 5; 41pp; English.

CC A method has been developed of producing a modified pertussis holotoxin,

CC involving analysis of the 3-dimensional form of the crystalline

CC holotoxin. The pertussis holotoxin modification process comprises:

CC (1) identification of at least one amino acid (aa) residue of the

CC holotoxin for modification by analysing the 3-dimensional form of the

CC crystalline holotoxin, in relation to known information of the protein

CC structure and function; (2) effecting mutagenesis (by removing or

CC replacing a nucleotide sequence encoding at least one (aa) of a tox

CC operon; and (3) expressing mutant tox box in a Bordetella organism to

CC produce the modified holotoxin. This method is used for modifying

CC pertussis holotoxin, by studying its 3-dimensional crystalline

CC structure. Modifying the holotoxin, alters its biological properties.

CC By analysing the 3-dimensional crystalline structure of the pertussis

CC holotoxin, functional (aa) which affect biological properties of the

CC pertussis holotoxin can be identified. This can be used to predict (aa)

CC which contribute to the toxicity of the holotoxin to produce

CC immunoprotective, genetically-detoxified analogues of pertussis

CC holotoxin. The present sequence represents an ADP-ribosylating toxin

CC B-subunit peptide used in the exemplification of the present

SQ Sequence 93 AA;

Query Match 100.0%; Score 39; DB 20; Length 93;

Best Local Similarity 100.0%; Pred. No. 2.4;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EVPGSOH 7

DB 41 evpgsqh 47

RESULT 15

ID AAW95226 standard; peptide; 93 AA.

AC AAW95226;
 XX
 DT 16-MAR-1999 (first entry)
 XX
 DE E. coli heat-labile toxin (LT) beta-subunit sequence.
 XX
 KW pertussis holotoxin; PT, modified; effector; toxicity; cell binding;
 KW enzyme; ADP-ribosylation; toxin; crystalline; X-ray crystallography;
 KW structural analysis; interacting site; mitogenicity; adjuvanticity;
 KW heat-labile; LT.
 XX
 OS Escherichia coli.
 XX
 PN US5856122-A.
 XX
 PD 05-JAN-1999.
 XX
 PF 22-AUG-1994; 94US-0292968.
 XX
 PR 22-AUG-1994; 94US-0292968.
 PR 24-AUG-1993; 93US-0110947.
 PR 31-MAY-1994; 94US-0251121.
 XX
 PA (UYAL-) UNIV ALBERTA.
 XX
 PI Armstrong GD, Cockle SA, Hazes B, Klein MH, Loosmore S;
 PI Oomen RP, Read RJ, Stein PE;
 XX
 DR WPI: 1999-105104/09.
 XX
 PT Modifications to e.g. enzymatic activity, mitogenicity and cell
 PT binding of pertussis holotoxin - by identifying interaction sites of
 PT a molecule with crystalline toxin and modifying the identified site
 XX
 PS Example 3; Fig 5; 40pp; English.
 XX
 CC The invention relates to methods of preparing a pertussis holotoxin (PT)
 CC having a modified biological activity. One method comprises identifying
 CC at least 1 site in a PT that interacts with a molecule that is capable of
 CC forming a complex with the holotoxin and which molecule is an effector
 CC molecule which is an adenine nucleotide and which site contributes to
 CC toxicity, cell binding or enzymatic activity of PT. The functional
 CC interacting site(s) are identified by analysing the three dimensional
 CC structure of crystalline PT, determined by X-ray crystallography. The
 CC identified interacting site(s) are modified to alter toxicity, cell
 CC binding or enzyme activity of the PT. The methods can be used to alter a
 CC biological activity such as toxicity, enzymatic activity, mitogenicity,
 CC cell binding and adjuvanticity of the PT. The three-dimensional structure
 CC of PT have functional and/or structural resemblance to other bacterial
 CC toxins such as diphtheria toxin (DT), Pseudomonas exotoxin A (ETA), the
 CC heat-labile toxin of E. coli (LT) and verotoxin-1 (VT). The present
 CC sequence represents the beta-subunit of LT toxin.
 XX
 SQ Sequence 93 AA;

Query Match 100.0%; Score 39; DB 20; Length 93;
 Best Local Similarity 100.0%; Pred. No. 2.4;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EYVGSQH 7
 |||||
 Db 41 evpsqsh 47

Search completed: July 3, 2002, 08:54:56
 Job time: 286 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 3, 2002, 08:52:05 ; Search time 34.88 Seconds
(without alignments)
4.902 Million cell updates/sec

Title: US-09-786-648-2

Perfect score: 39

Sequence: 1 EYPGSH 7

Scoring table:

BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 08
Maximum Match 1008

Listing first 45 summaries

Database :

Issued Patents AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	39	100.0	93	2	US-08-292-968-26 Sequence 26, Appl
2	39	100.0	93	2	US-08-467-974-26 Sequence 26, Appl
3	39	100.0	93	2	US-08-467-536-26 Sequence 26, Appl
4	39	100.0	93	3	US-08-467-976-26 Sequence 26, Appl
5	39	100.0	93	4	US-09-082-514-26 Sequence 26, Appl
6	39	100.0	102	3	US-08-952-337-5 Sequence 5, Appl
7	39	100.0	102	3	US-08-952-337-6 Sequence 6, Appl
8	39	100.0	103	2	US-08-472-171-2 Sequence 2, Appl
9	39	100.0	103	2	US-08-894-526-2 Sequence 2, Appl
10	39	100.0	103	4	US-09-013-047-2 Sequence 2, Appl
11	39	100.0	103	4	US-09-374-597-2 Sequence 2, Appl
12	39	100.0	103	5	US-09-191-852-21 Sequence 21, Appl
13	39	100.0	103	5	PCT-US95-13376-21 Sequence 21, Appl
14	39	100.0	123	3	US-08-952-337-1 Sequence 1, Appl
15	39	100.0	123	3	US-08-952-337-2 Sequence 2, Appl
16	39	100.0	124	2	US-08-747-410-2 Sequence 2, Appl
17	39	100.0	371	1	US-08-829-026A-6 Sequence 6, Appl
18	33	84.6	124	1	US-08-449-045C-4 Sequence 4, Appl
19	33	84.6	124	6	US-08-435-605A-12 Sequence 12, Appl
20	33	84.6	124	6	5223610-3 Patent No. 5223610
21	31	79.5	93	4	US-09-069-023-15 Sequence 15, Appl
22	31	79.5	101	4	US-09-031-962D-16 Sequence 16, Appl
23	31	79.5	101	4	US-09-031-962D-17 Sequence 17, Appl
24	31	79.5	219	4	US-09-069-023-12 Sequence 12, Appl
25	31	79.5	414	1	US-08-255-471-9 Sequence 9, Appl
26	31	79.5	459	6	5194375-6 Patent No. 5194375
27	30	76.9	262	6	5194375-4 Patent No. 5194375

ALIGNMENTS

28	30	76.9	329	2	US-08-781-802-8	Sequence 8, Appl
29	30	76.9	329	4	US-08-694-078-8	Sequence 8, Appl
30	30	76.9	329	4	US-09-058-260-8	Sequence 8, Appl
31	30	76.9	459	6	5194375-2	Patent No. 5194375
32	30	76.9	775	2	US-08-714-070A-1	Sequence 1, Appl
33	30	76.9	805	1	US-08-045-806-2	Sequence 2, Appl
34	30	76.9	805	1	US-08-366-051B-2	Sequence 2, Appl
35	30	76.9	855	2	US-09-027-337-2	Sequence 2, Appl
36	30	76.9	1285	2	US-08-540-406-6	Sequence 6, Appl
37	30	76.9	1285	3	US-08-656-055-6	Sequence 6, Appl
38	30	76.9	1285	4	US-08-954-668-6	Sequence 6, Appl
39	30	76.9	1285	5	PCT-US95-13233-6	Sequence 6, Appl
40	30	76.9	1286	4	US-09-268-140-3	Sequence 3, Appl
41	30	76.9	1299	4	US-08-460-900C-62	Sequence 62, Appl
42	30	76.9	1299	4	US-08-674-509B-48	Sequence 48, Appl
43	30	76.9	1299	4	US-08-954-698-48	Sequence 48, Appl
44	30	76.9	1810	5	PCT-US95-11684-4	Sequence 4, Appl
45	29	74.4	346	2	US-08-602-359A-34	Sequence 34, Appl

RESULT 1
US-08-292-968-26
Sequence 26, Application US/08292968
Patent No. 5856122
GENERAL INFORMATION:
APPLICANT: READ, Randy J.
APPLICANT: STEIN, Penelope E.
APPLICANT: COCKLE, Stephen A.
APPLICANT: OOMEN, Raymond P.
APPLICANT: LOOSMORE, Sheena
APPLICANT: KLEIN, Michel H.
APPLICANT: ARMSTRONG, Glen D.
TITLE OF INVENTION: MODIFICATION OF PERTUSSIS TOXIN
NUMBER OF SEQUENCES: 46
CORRESPONDENCE ADDRESS:
ADDRESSEE: Slim & McBurney
STREET: Suite 701, 330 University Avenue
City: Toronto
STATE: Ontario
COUNTRY: Canada
ZIP: M5G 1R7
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/292,968
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/110,947
FILING DATE: 24-AUG-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/251,121
FILING DATE: 31-MAY-1994
ATTORNEY/AGENT INFORMATION:
NAME: STEWART, Michael I.
REGISTRATION NUMBER: 24,973
REFERENCE/DOCKET NUMBER: 1038-388
TELECOMMUNICATION INFORMATION:
TELEPHONE: (416) 595-1155
TELEFAX: (416) 595-1163
INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 93 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear

US-08-292-968-26

Query Match 100.0%; Score 39; DB 2; Length 93;
Best Local Similarity 100.0%; Pred. No. 0.67;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EYVGSQH 7
|||||
DB 41 EYVGSQH 47

RESULT 2
US-08-467-974-26
; Sequence 26, Application US/08467974
; Patent No. 5965385
; GENERAL INFORMATION:
; APPLICANT: READ, Randy J.
; APPLICANT: STEIN, Penelope E.
; APPLICANT: COCKLE, Stephen A.
; APPLICANT: COOMEN, Raymond P.
; APPLICANT: LOOSMORE, Sheena
; APPLICANT: KLEIN, Michel H.
; APPLICANT: ARMSTRONG, Glen D.
; APPLICANT: HAZES, Bart
; TITLE OF INVENTION: MODIFICATION OF PERTUSSIS TOXIN
; NUMBER OF SEQUENCES: 46
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Slim & McBurney
; STREET: Suite 701, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/467,974
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/467,536
; FILING DATE: 06-JUN-1995
; APPLICATION NUMBER: US 08/292,968
; FILING DATE: 22-AUG-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/251,121
; FILING DATE: 31-MAY-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/110,947
; FILING DATE: 24-AUG-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: STEWART, Michael I.
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-454 MIS:vg
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; INFORMATION FOR SEQ ID NO: 26:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 93 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-467-974-26

Query Match 100.0%; Score 39; DB 2; Length 93;
Best Local Similarity 100.0%; Pred. No. 0.67;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EYVGSQH 7
|||||
DB 41 EYVGSQH 47

RESULT 3
US-08-467-536-26
; Sequence 26, Application US/08467536
; Patent No. 597304
; GENERAL INFORMATION:
; APPLICANT: READ, Randy J.
; APPLICANT: STEIN, Penelope E.
; APPLICANT: COCKLE, Stephen A.
; APPLICANT: COOMEN, Raymond P.
; APPLICANT: LOOSMORE, Sheena
; APPLICANT: KLEIN, Michel H.
; APPLICANT: ARMSTRONG, Glen D.
; APPLICANT: HAZES, Bart
; TITLE OF INVENTION: MODIFICATION OF PERTUSSIS TOXIN
; NUMBER OF SEQUENCES: 46
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Slim & McBurney
; STREET: Suite 701, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/467,536
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/292,968
; FILING DATE: 22-AUG-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/251,121
; FILING DATE: 31-MAY-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/110,947
; FILING DATE: 24-AUG-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: STEWART, Michael I.
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-455 MIS:vg
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; INFORMATION FOR SEQ ID NO: 26:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 93 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-467-536-26

Query Match 100.0%; Score 39; DB 2; Length 93;
Best Local Similarity 100.0%; Pred. No. 0.67;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EYVGSQH 7
|||||
DB 41 EYVGSQH 47

RESULT 4
US-08-467-976-26

Sequence 26, Application US/08467976
Patent No. 6018022
GENERAL INFORMATION:
APPLICANT: READ, Randy J.
APPLICANT: STEIN, Penelope E.
APPLICANT: COCKLE, Stephen A.
APPLICANT: COOMEN, Raymond P.
APPLICANT: LOOSMORE, Sheena
APPLICANT: KLEIN, Michel H.
APPLICANT: ARMSTRONG, Glen D.
TITLE OF INVENTION: MODIFICATION OF PERTUSSIS TOXIN
NUMBER OF SEQUENCES: 46
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sim & McBurney
STREET: Suite 701, 330 University Avenue
CITY: Toronto
STATE: Ontario
COUNTRY: Canada
ZIP: M5G 1R7
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/467,976
FILING DATE: 06-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/292,968
FILING DATE: 22-AUG-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/251,121
FILING DATE: 31-MAY-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/110,947
FILING DATE: 24-AUG-1993
ATTORNEY/AGENT INFORMATION:
NAME: STEWART, Michael I.
REGISTRATION NUMBER: 24,973
REFERENCE/DOCKET NUMBER: 1038-453 MIS:vg
TELECOMMUNICATION INFORMATION:
TELEPHONE: (416) 595-1155
TELEFAX: (416) 595-1163
INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 93 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-467-976-26

Query Match 100.0%; Score 39; DB 3; Length 93;
Best Local Similarity 100.0%; Pred. No. 0.67;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EYVGSQH 7
|||||
DB 41 EYVGSQH 47

RESULT 5
US-09-082-514-26
Sequence 26, Application US/09082514
Patent No. 6168928
GENERAL INFORMATION:
APPLICANT: READ, Randy J.
APPLICANT: STEIN, Penelope E.
APPLICANT: COCKLE, Stephen A.
APPLICANT: COOMEN, Raymond P.
APPLICANT: KLEIN, Michel H.

APPLICANT: ARMSTRONG, Glen D.
APPLICANT: HAZES, Bart
TITLE OF INVENTION: MODIFICATION OF PERTUSSIS TOXIN
NUMBER OF SEQUENCES: 46
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sim & McBurney
STREET: 6th Floor, 330 University Avenue
CITY: Toronto
STATE: Ontario
COUNTRY: Canada
ZIP: M5G 1R7
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/082,514
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/292,968
FILING DATE: 24-AUG-1994
ATTORNEY/AGENT INFORMATION:
NAME: STEWART, Michael I.
REGISTRATION NUMBER: 24,973
REFERENCE/DOCKET NUMBER: 1038-810
TELECOMMUNICATION INFORMATION:
TELEPHONE: (416) 595-1155
TELEFAX: (416) 595-1163
INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 93 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-09-082-514-26

Query Match 100.0%; Score 39; DB 4; Length 93;
Best Local Similarity 100.0%; Pred. No. 0.67;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EYVGSQH 7
|||||
DB 41 EYVGSQH 47

RESULT 6
US-08-952-337-5
Sequence 5, Application US/08952337
Patent No. 6019973
GENERAL INFORMATION:
APPLICANT: Holmgren, Jan
APPLICANT: Iebens, Michael R.
TITLE OF INVENTION: HYBRID MOLECULES BETWEEN HEAT-LABILE
TITLE OF INVENTION: ENTEROTOXIN AND CHOLERA TOXIN B SUBUNITS
FILE REFERENCE: 3846/0D758
CURRENT APPLICATION NUMBER: US/08/952,337
CURRENT FILING DATE: 1998-01-05
EARLIER APPLICATION NUMBER: PCT/SE96/00570
EARLIER FILING DATE: 1996-05-02
EARLIER APPLICATION NUMBER: SE 9501682-0
EARLIER FILING DATE: 1995-05-05
NUMBER OF SEQ ID NOS: 6
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 5
LENGTH: 102
TYPE: PRT
ORGANISM: Vibrio cholerae
US-08-952-337-5

Query Match 100.0%; Score 39; DB 3; Length 102;
Best Local Similarity 100.0%; Pred. No. 0.73;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EYPSGSH 7
|||||||
DB 50 EYPSGSH 56

RESULT 7

US-08-952-337-6
; Sequence 6, Application US/08952337
; Patent No. 6019973
; GENERAL INFORMATION:
; APPLICANT: Holmgren, Jan
; APPLICANT: Lebens, Michael R.
; TITLE OF INVENTION: HYBRID MOLECULES BETWEEN HEAT-LABILE
; TITLE OF INVENTION: ENTEROTOXIN AND CHOLERA TOXIN B SUBUNITS
; FILE REFERENCE: 3846/OD758
; CURRENT FILING DATE: 1998-01-05
; EARLIER FILING DATE: 1996-05-02
; EARLIER FILING DATE: 1996-05-02
; EARLIER APPLICATION NUMBER: SE 9501682-0
; EARLIER FILING DATE: 1995-05-05
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FASTSEQ for Windows Version 3.0
; SEQ ID NO 6
; LENGTH: 102
; TYPE: PRT
; ORGANISM: Escherichia coli
US-08-952-337-6

Query Match 100.0%; Score 39; DB 3; Length 102;
Best Local Similarity 100.0%; Pred. No. 0.73;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EYPSGSH 7
|||||||
DB 50 EYPSGSH 56

RESULT 8

US-08-472-171-2
; Sequence 2, Application US/08472171
; Patent No. 5932714
; GENERAL INFORMATION:
; APPLICANT: Loosmore, Sheena M.
; APPLICANT: Yacoub, Reza K.
; APPLICANT: Zealey, Gavin R.
; APPLICANT: Klein, Michel H.
; TITLE OF INVENTION: Expression Of Gene Products From
; TITLE OF INVENTION: Genetically Manipulated Strains Of Bordetella
; NUMBER OF SEQUENCES: 56
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: 330 University Avenue, Suite 701
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/393,334

; FILING DATE: 23-FEB-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Stewart, Michael I.
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-507 MIS.vg
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 416-595-1155
; TELEFAX: 416-595-1163
; TELEX: 065-24567 Simbas
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 103 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-472-171-2

Query Match 100.0%; Score 39; DB 2; Length 103;
Best Local Similarity 100.0%; Pred. No. 0.74;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EYPSGSH 7
|||||||
DB 51 EYPSGSH 57

RESULT 9

US-08-894-526-2
; Sequence 2, Application US/08894526
; Patent No. 5942418
; GENERAL INFORMATION:
; APPLICANT: Loosmore, Sheena M
; APPLICANT: Yacoub, Reza K
; APPLICANT: Zealey, Gavin R
; APPLICANT: Klein, Michel H
; TITLE OF INVENTION: EXPRESSION OF GENE PRODUCTS FROM
; TITLE OF INVENTION: GENETICALLY MANIPULATED STRAINS OF BORDETELLA
; NUMBER OF SEQUENCES: 57
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: 6th Floor, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/894,526
; FILING DATE: 01-DEC-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Stewart, Michael I
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-724 MIS.jb
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 103 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-894-526-2

Query Match 100.0%; Score 39; DB 2; Length 103;
Best Local Similarity 100.0%; Pred. No. 0.74;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EVPGSOH 7
|||||
Db 51 EVPGSOH 57

RESULT 10

US-09-013-047-2
; Sequence 2, Application US/09013047
; Patent No. 5998168
; GENERAL INFORMATION:
; APPLICANT: Loosmore, Sheena M.
; APPLICANT: Yaccob, Reza K.
; APPLICANT: Zealey, Gavin R.
; APPLICANT: Klein, Michel H.
; TITLE OF INVENTION: Expression Of Gene Products From
; TITLE OF INVENTION: Genetically Manipulated Strains Of Bordetella
; NUMBER OF SEQUENCES: 56
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Sim & Mcburney
; STREET: 330 University Avenue, 6th Floor
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/013,047
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/472,171
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/393,334
; FILING DATE: 23-FEB-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Stewart, Michael I.
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-507 MTS:vg
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 416-595-1155
; TELEFAX: 065-24567 Simbas
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 103 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-013-047-2

Query Match 100.0%; Score 39; DB 2; Length 103;
Best Local Similarity 100.0%; Pred. No. 0.74;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EVPGSOH 7
|||||
Db 51 EVPGSOH 57

RESULT 11
US-09-374-597-2
; Sequence 2, Application US/09374597
; Patent No. 6140082
; GENERAL INFORMATION:
; APPLICANT: Loosmore, Sheena M.

APPLICANT: Yaccob, Reza K.
APPLICANT: Zealey, Gavin R.
APPLICANT: Klein, Michel H.
TITLE OF INVENTION: Expression Of Gene Products From
TITLE OF INVENTION: Genetically Manipulated Strains Of Bordetella
NUMBER OF SEQUENCES: 56
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Sim & Mcburney
STREET: 330 University Avenue, 6th Floor
CITY: Toronto
STATE: Ontario
COUNTRY: Canada
ZIP: M5G 1R7

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/374,597
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/393,334
FILING DATE: FEBRUARY 23, 1995
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Stewart, Michael I
REGISTRATION NUMBER: 24973
REFERENCE/DOCKET NUMBER: 1038-964
TELECOMMUNICATION INFORMATION:
TELEPHONE: 416-595-1155
TELEFAX: 416-595-1163

INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 103 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-374-597-2

Query Match 100.0%; Score 39; DB 4; Length 103;
Best Local Similarity 100.0%; Pred. No. 0.74;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EVPGSOH 7
|||||
Db 51 EVPGSOH 57

RESULT 12
US-09-191-852-21
; Sequence 21, Application US/09191852
; Patent No. 6194560
; GENERAL INFORMATION:
; APPLICANT: Charles J. Arntzen, Hugh S. Mason, and Tariq A. Haq
; TITLE OF INVENTION: ORAL IMMUNIZATION WITH TRANSGENIC PLANTS
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Fulbright & Jaworski L.L.P.
; STREET: 1301 McKinney, Suite 5100
; CITY: Houston
; STATE: TX
; COUNTRY: USA
; ZIP: 77010

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/191,852
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/13376
FILING DATE: 24-OCT-1995
APPLICATION NUMBER: 08/817,906
FILING DATE: 04-AUG-1997
ATTORNEY/AGENT INFORMATION:
NAME: Fox, David L.
REGISTRATION NUMBER: 40,612
REFERENCE/DOCKET NUMBER: P015900S1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 713-651-5151
TELEFAX: 713-651-5246
INFORMATION FOR SEQ ID NO: 21:
SEQUENCE CHARACTERISTICS:
LENGTH: 103 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-09-191-852-21

Query Match 100.0%; Score 39; DB 4; Length 103;
Best Local Similarity 100.0%; Pred. No. 0.74;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EYVGSOH 7
|||||||
DB 51 EYVGSOH 57

RESULT 13
PCT-US95-13376-21
Sequence 21, Application PC/TUS9513376
GENERAL INFORMATION:
APPLICANT: The Texas A&M University System
APPLICANT: 310 Wisenbaker
APPLICANT: College Station, Texas 77843-3369
TITLE OF INVENTION: ORAL IMMUNIZATION WITH TRANSGENIC PLANTS
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pravel, Hewitt, Kimball & Krieger
STREET: 1177 West Loop South, 10th Floor
CITY: Houston
STATE: TX
COUNTRY: USA
ZIP: 77027-9095
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/13376
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/328,716
FILING DATE: 24-OCT-1994
ATTORNEY/AGENT INFORMATION:
NAME: Jones, John W.
REGISTRATION NUMBER: 31,380
REFERENCE/DOCKET NUMBER: 36170/3P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 713-850-0909
TELEFAX: 713-850-0165
INFORMATION FOR SEQ ID NO: 21:
SEQUENCE CHARACTERISTICS:
LENGTH: 103 amino acids
TYPE: amino acid
TOPOLOGY: linear
PCT-US95-13376-21

Query Match 100.0%; Score 39; DB 5; Length 103;
Best Local Similarity 100.0%; Pred. No. 0.74;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EYVGSOH 7
|||||||
DB 51 EYVGSOH 57

RESULT 14
US-08-952-337-1
Sequence 1, Application US/08952337
Patent No. 6019973
GENERAL INFORMATION:
APPLICANT: Holmgren, Jan
APPLICANT: Lebens, Michael R.
TITLE OF INVENTION: HYBRID MOLECULES BETWEEN HEAT-LABILE
TITLE OF INVENTION: ENTEROTOXIN AND CHOLERA TOXIN B SUBUNITS
FILE REFERENCE: 3846/OD758
CURRENT APPLICATION NUMBER: US/08/952,337
CURRENT FILING DATE: 1998-01-05
EARLIER APPLICATION NUMBER: PCT/SE96/00570
EARLIER FILING DATE: 1996-05-02
EARLIER APPLICATION NUMBER: SE 9501682-0
EARLIER FILING DATE: 1995-05-05
NUMBER OF SEQ ID NOS: 6
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 1
LENGTH: 123
TYPE: PRT
ORGANISM: Vibrio cholerae
US-08-952-337-1

Query Match 100.0%; Score 39; DB 3; Length 123;
Best Local Similarity 100.0%; Pred. No. 0.9;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EYVGSOH 7
|||||||
DB 71 EYVGSOH 77

RESULT 15
US-08-952-337-2
Sequence 2, Application US/08952337
Patent No. 6019973
GENERAL INFORMATION:
APPLICANT: Holmgren, Jan
APPLICANT: Lebens, Michael R.
TITLE OF INVENTION: ENTEROTOXIN AND CHOLERA TOXIN B SUBUNITS
FILE REFERENCE: 3846/OD758
CURRENT APPLICATION NUMBER: US/08/952,337
CURRENT FILING DATE: 1998-01-05
EARLIER APPLICATION NUMBER: PCT/SE96/00570
EARLIER FILING DATE: 1996-05-02
EARLIER APPLICATION NUMBER: SE 9501682-0
EARLIER FILING DATE: 1995-05-05
NUMBER OF SEQ ID NOS: 6
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 2
LENGTH: 123
TYPE: PRT
ORGANISM: Escherichia coli
US-08-952-337-2

Query Match 100.0%; Score 39; DB 3; Length 123;
Best Local Similarity 100.0%; Pred. No. 0.9;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	EVPGSQH	7
Db	71	EVPGSQH	77

Search completed: July 3, 2002, 08:55:38
Job time: 213 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 3, 2002, 08:53:00 ; Search time 401.04 Seconds
(without alignments)
6.144 Million cell updates/sec

Title: US-09-786-648-2

Perfect score: 39

Sequence: 1 EVPGSQH 7

Scoring table:

BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 3502263 seqs, 351980561 residues

Total number of hits satisfying chosen parameters: 3502263

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : Listing first 45 summaries

Pending_Patents_AA_Main:*

1: /cgn2_6/ptodata/2/paa/PCTUS.COMB.pep.*
2: /cgn2_6/ptodata/2/paa/US06.COMB.pep.*
3: /cgn2_6/ptodata/2/paa/US07.COMB.pep.*
4: /cgn2_6/ptodata/2/paa/US08.COMB.pep.*
5: /cgn2_6/ptodata/2/paa/US081.COMB.pep.*
6: /cgn2_6/ptodata/2/paa/US082.COMB.pep.*
7: /cgn2_6/ptodata/2/paa/US083.COMB.pep.*
8: /cgn2_6/ptodata/2/paa/US084.COMB.pep.*
9: /cgn2_6/ptodata/2/paa/US085.COMB.pep.*
10: /cgn2_6/ptodata/2/paa/US086.COMB.pep.*
11: /cgn2_6/ptodata/2/paa/US087.COMB.pep.*
12: /cgn2_6/ptodata/2/paa/US088.COMB.pep.*
13: /cgn2_6/ptodata/2/paa/US089.COMB.pep.*
14: /cgn2_6/ptodata/2/paa/US090.COMB.pep.*
15: /cgn2_6/ptodata/2/paa/US091.COMB.pep.*
16: /cgn2_6/ptodata/2/paa/US092.COMB.pep.*
17: /cgn2_6/ptodata/2/paa/US093.COMB.pep.*
18: /cgn2_6/ptodata/2/paa/US094.COMB.pep.*
19: /cgn2_6/ptodata/2/paa/US095.COMB.pep.*
20: /cgn2_6/ptodata/2/paa/US096.COMB.pep.*
21: /cgn2_6/ptodata/2/paa/US097.COMB.pep.*
22: /cgn2_6/ptodata/2/paa/US098.COMB.pep.*
23: /cgn2_6/ptodata/2/paa/US099.COMB.pep.*
24: /cgn2_6/ptodata/2/paa/US100.COMB.pep.*
25: /cgn2_6/ptodata/2/paa/US101.COMB.pep.*
26: /cgn2_6/ptodata/2/paa/US060.COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	39	100.0	7	21	US-09-786-648-2
2	39	100.0	12	21	US-09-786-648-3
3	39	100.0	15	11	US-08-732-371-1
4	39	100.0	15	11	US-08-732-371A-1
5	39	100.0	21	21	US-09-786-648-4
6	39	100.0	21	21	US-09-786-648-5
7	39	100.0	93	5	US-08-110-947-10

8	39	100.0	93	5	US-08-110-947A-26	Sequence 26, Appl
9	39	100.0	93	6	US-08-251-121-26	Sequence 26, Appl
10	39	100.0	103	7	US-08-393-334-2	Sequence 2, Appl1
11	39	100.0	103	7	US-08-782-832-15	Sequence 15, Appl
12	39	100.0	103	12	US-08-817-906-21	Sequence 21, Appl
13	39	100.0	103	22	US-09-836-433-14	Sequence 14, Appl
14	39	100.0	116	22	US-09-836-433-20	Sequence 20, Appl
15	39	100.0	119	22	US-09-836-433-22	Sequence 22, Appl
16	39	100.0	124	1	PCT-US99-30747-55	Sequence 55, Appl
17	39	100.0	124	1	PCT-US99-30747-57	Sequence 55, Appl
18	39	100.0	124	18	US-09-470-124-55	Sequence 55, Appl
19	39	100.0	124	18	US-09-470-124-57	Sequence 57, Appl
20	39	100.0	138	13	US-08-914-479-2	Sequence 2, Appl1
21	39	100.0	138	13	US-08-914-479A-2	Sequence 2, Appl1
22	39	100.0	133	21	US-09-756-983-15	Sequence 15, Appl
23	39	100.0	351	21	US-09-756-983-18	Sequence 18, Appl
24	39	100.0	364	21	US-09-756-983-22	Sequence 22, Appl
25	39	100.0	371	5	US-08-150-305A-3	Sequence 3, Appl1
26	39	100.0	371	11	US-08-784-218-6	Sequence 6, Appl1
27	39	100.0	371	12	US-08-829-026-5	Sequence 5, Appl1
28	39	89.7	142	26	US-60-361-742-1547	Sequence 1547, Ap
29	39	87.2	41	20	US-09-688-051-2967	Sequence 2967, Ap
30	34	87.2	84	20	US-09-617-682A-5651	Sequence 5651, Ap
31	34	87.2	124	18	US-09-417-507-29272	Sequence 29272, A
32	34	87.2	304	1	PCT-US01-08631-32859	Sequence 32859, A
33	34	87.2	346	1	PCT-US01-08631-32459	Sequence 32459, A
34	34	87.2	578	1	PCT-US01-08631-32551	Sequence 32551, A
35	34	87.2	847	1	PCT-US01-08631-54035	Sequence 54035, A
36	33	84.6	51	22	US-09-866-066-35	Sequence 35, Appl
37	33	84.6	124	21	US-09-760-234-7	Sequence 7, Appl1
38	33	84.6	131	26	US-60-171-481-1502	Sequence 1502, Ap
39	33	84.6	318	1	PCT-US99-17130-318	Sequence 318, App
40	33	84.6	318	18	US-09-489-847-3328	Sequence 328, App
41	33	84.6	379	18	US-09-402-532-1	Sequence 1, Appl1
42	33	84.6	379	18	US-09-402-532-3	Sequence 3, Appl1
43	33	84.6	382	1	PCT-US01-08582-3	Sequence 3, Appl1
44	33	84.6	382	1	PCT-US01-08582-4	Sequence 4, Appl1
45	33	84.6	382	18	US-09-402-532-4	Sequence 4, Appl1

ALIGNMENTS

RESULT 1
US-09-786-648-2
; Sequence 2, Application US/09786648
; GENERAL INFORMATION:
; APPLICANT: Williams, Neil Andrew
; TITLE OF INVENTION: Peptide Fragments of Cholera Toxin B or Enterotoxin B as Vacci
; FILE REFERENCE: 7438
; CURRENT APPLICATION NUMBER: US/09/786, 648
; PRIOR APPLICATION NUMBER: 2001-03-07
; PRIOR FILING DATE: 1999-09-07
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: MS DOS
; SEQ ID NO 2
; LENGTH: 7
; TYPE: PRT
; ORGANISM: E. coli
; FEATURE:
; LOCATION: 51...57
; OTHER INFORMATION: Isolated or synthetic E. coli
; OTHER INFORMATION: human variant E. coli
US-09-786-648-2

Query Match 100.0%; Score 39; DB 21; Length 7;
Best Local Similarity 100.0%; Pred. No. 3.2e+06;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EYVGSQH 7
| | | | |
DB 1 EYVGSQH 7

RESULT 2

US-09-786-648-3
; Sequence 3, Application US/09786648
; GENERAL INFORMATION:
; APPLICANT: Williams, Neil Andrew
; APPLICANT: Hirst, Timothy Raymond
; TITLE OF INVENTION: Peptide Fragments of Cholera Toxin B or Enterotoxin B as Vaccine
; TITLE OF INVENTION: Adjuvants
; FILE REFERENCE: 7438
; CURRENT APPLICATION NUMBER: US/09/786,648
; CURRENT FILING DATE: 2001-03-07
; PRIOR APPLICATION NUMBER: PCT/GB99/02970
; PRIOR FILING DATE: 1999-09-07
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: MS DOS
; SEQ ID NO 3
; LENGTH: 12
; TYPE: PRT
; ORGANISM: E. coli
; FEATURE:
; LOCATION: 50...61
; OTHER INFORMATION: Isolated or synthetic EtXB beta4-alpha2 loop fragment derivable f
US-09-786-648-3

Query Match 100.0%; Score 39; DB 21; Length 12;
Best Local Similarity 100.0%; Pred. No. 2.4;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EYVGSQH 7
| | | | |
DB 2 EYVGSQH 8

RESULT 3

US-08-732-371-1
; Sequence 1, Application US/08732371
; GENERAL INFORMATION:
; APPLICANT: MIRELMAN, David
; APPLICANT: MARKS, Robert S.
; APPLICANT: SELA, Michael
; TITLE OF INVENTION: VACCINES FOR ORAL IMMUNIZATION AGAINST
; TITLE OF INVENTION: INFECTING AGENTS
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWDY AND NEIMARK
; STREET: 419 Seventh Street, N.W., Suite 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/732,371
; FILING DATE: 09-JAN-1997
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: IL 109519
; FILING DATE: 03-MAY-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: YUN, Allen C.
; REGISTRATION NUMBER: 37,971
; REFERENCE/DOCKET NUMBER: MIRELMAN-3

TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-628-5197
; TELEFAX: 202-737-3528
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-732-371-1

Query Match 100.0%; Score 39; DB 11; Length 15;
Best Local Similarity 100.0%; Pred. No. 3.1;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EYVGSQH 7
| | | | |
DB 2 EYVGSQH 8

RESULT 4

US-08-732-371A-1
; Sequence 1, Application US/08732371A
; GENERAL INFORMATION:
; APPLICANT: MIRELMAN, David
; APPLICANT: MARKS, Robert S.
; APPLICANT: SELA, Michael
; TITLE OF INVENTION: VACCINES FOR ORAL IMMUNIZATION AGAINST
; TITLE OF INVENTION: INFECTING AGENTS
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWDY AND NEIMARK
; STREET: 419 Seventh Street, N.W., Suite 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; OPERATING SYSTEM: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/732,371A
; FILING DATE: 09-JAN-1997
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: IL 109519
; FILING DATE: 03-MAY-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: YUN, Allen C.
; REGISTRATION NUMBER: 37,971
; REFERENCE/DOCKET NUMBER: MIRELMAN-3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-628-5197
; TELEFAX: 202-737-3528
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-732-371A-1

Query Match 100.0%; Score 39; DB 11; Length 15;
Best Local Similarity 100.0%; Pred. No. 3.1;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EYVGSQH 7
| | | | |
DB 2 EYVGSQH 8

RESULT 5
US-09-786-648-4
; Sequence 4, Application US/09786648
; GENERAL INFORMATION:
; APPLICANT: Williams, Neil Andrew
; APPLICANT: Hirst, Timothy Raymond
; TITLE OF INVENTION: Peptide Fragments of Cholera Toxin B or Enterotoxin B as Vaccine
; FILE REFERENCE: 7438
; CURRENT APPLICATION NUMBER: US/09/786,648
; CURRENT FILING DATE: 2001-03-07
; PRIOR APPLICATION NUMBER: PCT/GB99/02970
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: MS DOS
; SEQ ID NO 4
; LENGTH: 21
; TYPE: PRT
; ORGANISM: E. coli
; FEATURE:
; LOCATION: 45...65
; OTHER INFORMATION: isolated or synthetic EtxB beta4-alpha2 loop fragment derivable
; OTHER INFORMATION: human variant E. coli
US-09-786-648-4

Query Match 100.0%; Score 39; DB 21; Length 21;
Best Local Similarity 100.0%; Pred. No. 4.4;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EYVGSQH 7
|||||
DB 7 EYVGSQH 13

RESULT 6
US-09-786-648-5
; Sequence 5, Application US/09786648
; GENERAL INFORMATION:
; APPLICANT: Williams, Neil Andrew
; APPLICANT: Hirst, Timothy Raymond
; TITLE OF INVENTION: Peptide Fragments of Cholera Toxin B or Enterotoxin B as Vaccine
; FILE REFERENCE: 7438
; CURRENT APPLICATION NUMBER: US/09/786,648
; CURRENT FILING DATE: 2001-03-07
; PRIOR APPLICATION NUMBER: PCT/GB99/02970
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: MS DOS
; SEQ ID NO 5
; LENGTH: 21
; TYPE: PRT
; ORGANISM: E. coli
; FEATURE:
; LOCATION: 45...65
; OTHER INFORMATION: isolated or synthetic EtxB beta4-alpha2 loop fragment derivable
; OTHER INFORMATION: porcine E. coli
US-09-786-648-5

Query Match 100.0%; Score 39; DB 21; Length 21;
Best Local Similarity 100.0%; Pred. No. 4.4;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EYVGSQH 7
|||||
DB 7 EYVGSQH 13

RESULT 7

US-08-110-947-10
; Sequence 10, Application US/08110947
; GENERAL INFORMATION:
; APPLICANT: READ, Randy J
; APPLICANT: STEIN, Penelope E
; APPLICANT: COCKLE, Stephen A
; APPLICANT: OOMEN, Raymond P
; APPLICANT: KLEIN, Michel H
; TITLE OF INVENTION: MODIFICATION OF PERTUSSIS TOXIN
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Shoemaker and Mattare, Ltd.
; STREET: Suite 1203 Crystal Plaza I, 2001 Jefferson
; STREET: Davis Hwy.
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202-0286
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/110,947
; FILING DATE: 24-AUG-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: FALLOW, Charles W
; REGISTRATION NUMBER: 28,946
; REFERENCE/DOCKET NUMBER: 1038-303 MTS:jb
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 415-0810
; TELEFAX: (703) 415-0813
; TELEX: 89-9456 LUKPAT
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 93 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-110-947-10

Query Match 100.0%; Score 39; DB 5; Length 93;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EYVGSQH 7
|||||
DB 41 EYVGSQH 47

RESULT 8
US-08-110-947A-26
; Sequence 26, Application US/08110947A
; GENERAL INFORMATION:
; APPLICANT: READ, Randy J
; APPLICANT: STEIN, Penelope E
; APPLICANT: COCKLE, Stephen A
; APPLICANT: OOMEN, Raymond P
; APPLICANT: KLEIN, Michel H
; TITLE OF INVENTION: MODIFICATION OF PERTUSSIS TOXIN
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Shoemaker and Mattare, Ltd.
; STREET: Suite 1203 Crystal Plaza I, 2001 Jefferson
; STREET: Davis Highway
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202-0286
; COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/110,947A
FILING DATE: 24-AUG-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: FALLOM, Charles W
REGISTRATION NUMBER: 28,946
REFERENCE/DOCKET NUMBER: 1038-303 MIS-Jb
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 415-0810
TELEFAX: (703) 415-0813
TELEX: 89-9456 LUKPAR
INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 93 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-110-947A-26

Query Match 100.0%; Score 39; DB 5; Length 93;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EVPGSOH 7
Db 41 EVPGSOH 47

RESULT 9

US-08-251-121-26

Sequence 26, Application US/08251121
GENERAL INFORMATION:

APPLICANT: READ, Randy J.
APPLICANT: STEIN, Penelope E.
APPLICANT: COCKLE, Stephen A.
APPLICANT: COOMEN, Raymond P.
APPLICANT: KLEIN, Michael H.
APPLICANT: ARMSTRONG, Glen D.
APPLICANT: HAZES, Bart
TITLE OF INVENTION: MODIFICATION OF PERTUSSIS TOXIN
NUMBER OF SEQUENCES: 46
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sim & McBurney
STREET: Suite 701, 330 University Avenue
CITY: Toronto
STATE: Ontario
COUNTRY: Canada
ZIP: M5G 1R7

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/251,121
FILING DATE: 31-MAY-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/110,947
FILING DATE: 24-AUG-1993
ATTORNEY/AGENT INFORMATION:
NAME: STEWART, Michael I.
REGISTRATION NUMBER: 24,973
REFERENCE/DOCKET NUMBER: 1038-335
TELECOMMUNICATION INFORMATION:
TELEPHONE: (416) 595-1155
TELEFAX: (416) 595-1163

INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 93 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-251-121-26

Query Match 100.0%; Score 39; DB 6; Length 93;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EVPGSOH 7
Db 41 EVPGSOH 47

RESULT 10

US-08-393-334-2

Sequence 2, Application US/08393334
GENERAL INFORMATION:

APPLICANT: Loosmore, Sheena M.
APPLICANT: Yacoub, Reza K.
APPLICANT: Zealey, Gavin R.
APPLICANT: Klein, Michael H.
TITLE OF INVENTION: Expression Of Gene Products From
Genetically Manipulated Strains Of Bordetella
NUMBER OF SEQUENCES: 56
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sim & McBurney
STREET: 330 University Avenue, Suite 701
CITY: Toronto
STATE: Ontario
COUNTRY: Canada
ZIP: M5G 1R7

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/393,334
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Stewart, Michael I
REGISTRATION NUMBER: 24973
REFERENCE/DOCKET NUMBER: 1038-417
TELECOMMUNICATION INFORMATION:
TELEPHONE: 416-595-1163
TELEFAX: 416-595-1163
TELEX: 065-24567 Simbas
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 103 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-393-334-2

Query Match 100.0%; Score 39; DB 7; Length 103;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EVPGSOH 7
Db 51 EVPGSOH 57

RESULT 11
US-08-782-832-15

Sequence 15, Application US/08782832
GENERAL INFORMATION:
APPLICANT: Arnzen, Charles J.
APPLICANT: Mason, Hugh S.
TITLE OF INVENTION: PRODUCTION OF AN ORALLY IMMUNOGENIC
NUMBER OF SEQUENCES: 17
TITLE OF INVENTION: BACTERIAL PROTEIN IN TRANSGENIC PLANTS
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pravel, Hewitt, Kimball & Krieger
STREET: 1177 West Loop South, 10th floor
CITY: Houston
STATE: TX
COUNTRY: USA
ZIP: 77027-9095
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/782,832
FILING DATE: 13-JAN-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/328,716
FILING DATE: 24-OCT-1994
ATTORNEY/AGENT INFORMATION:
NAME: Kettelberger, Denise M.
REGISTRATION NUMBER: 33,924
REFERENCE/DOCKET NUMBER: 36170/3
TELEPHONE: 713-850-0909
TELEFAX: 713-850-0165
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 103 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-782-832-15

Query Match 100.0%; Score 39; DB 11; Length 103;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EVPGSOH 7
|||||||
Db 51 EVPGSOH 57

RESULT 12
US-08-817-906-21
Sequence 21, Application US/08817906
GENERAL INFORMATION:
APPLICANT: Charles J. Arnzen, Hugh S. Mason, John D. Clements,
APPLICANT: Haq, Tariq A.
TITLE OF INVENTION: ORAL IMMUNIZATION WITH TRANSGENIC PLANTS
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fulbright & Jaworski L.L.P.
STREET: 1301 McKinney, Suite 5100
CITY: Houston
STATE: TX
COUNTRY: USA
ZIP: 77010
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/817,906

FILING DATE: 08/04/97
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/13376
FILING DATE: 24-OCT-1995
ATTORNEY/AGENT INFORMATION:
NAME: Fox, David L.
REGISTRATION NUMBER: 40,612
REFERENCE/DOCKET NUMBER: 36170/3P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 713-651-5151
TELEFAX: 713-651-5246
INFORMATION FOR SEQ ID NO: 21:
SEQUENCE CHARACTERISTICS:
LENGTH: 103 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-817-906-21

Query Match 100.0%; Score 39; DB 12; Length 103;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EVPGSOH 7
|||||||
Db 51 EVPGSOH 57

RESULT 13
US-09-836-433-14
Sequence 14, Application US/09836433
GENERAL INFORMATION:
APPLICANT: Yuki, Yoshikazu
APPLICANT: Uda, Shigezo
TITLE OF INVENTION: HYBRID PROTEINS FOR AUTOIMMUNE DISEASE
FILE REFERENCE: Not Assigned
CURRENT APPLICATION NUMBER: US/09/836,433
CURRENT FILING DATE: 2001-04-16
NUMBER OF SEQ ID NOS: 49
SOFTWARE: Patentin version 3.0
SEQ ID NO 14
LENGTH: 103
TYPE: PRT
ORGANISM: Vibrio cholerae
US-09-836-433-14

Query Match 100.0%; Score 39; DB 22; Length 103;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EVPGSOH 7
|||||||
Db 51 EVPGSOH 57

RESULT 14
US-09-836-433-20
Sequence 20, Application US/09836433
GENERAL INFORMATION:
APPLICANT: Yuki, Yoshikazu
APPLICANT: Uda, Shigezo
TITLE OF INVENTION: HYBRID PROTEINS FOR AUTOIMMUNE DISEASE
FILE REFERENCE: Not Assigned
CURRENT APPLICATION NUMBER: US/09/836,433
CURRENT FILING DATE: 2001-04-16
NUMBER OF SEQ ID NOS: 49
SOFTWARE: Patentin version 3.0
SEQ ID NO 20
LENGTH: 116
TYPE: PRT
ORGANISM: synthetic construct

US-09-836-433-20

Query Match 100.0%; Score 39; DB 22; Length 116;
 Best Local Similarity 100.0%; Pred. No. 26;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EVPGSOH 7
 |||||
 DB 51 EVPGSOH 57

RESULT 15

US-09-836-433-22
 ; Sequence 22, Application US/09836433
 ; GENERAL INFORMATION:
 ; APPLICANT: Yuki, Yoshikazu
 ; APPLICANT: Udaoka, Shigezo
 ; TITLE OF INVENTION: HYBRID PROTEINS FOR AUTOIMMUNE DISEASE
 ; FILE REFERENCE: Not Assigned
 ; CURRENT APPLICATION NUMBER: US/09/836,433
 ; CURRENT FILING DATE: 2001-04-16
 ; NUMBER OF SEQ ID NOS: 49
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 22
 ; LENGTH: 119
 ; TYPE: PRT
 ; ORGANISM: synthetic construct
 US-09-836-433-22

Query Match 100.0%; Score 39; DB 22; Length 119;
 Best Local Similarity 100.0%; Pred. No. 27;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EVPGSOH 7
 |||||
 DB 51 EVPGSOH 57

Search completed: July 3, 2002, 09:03:20
 Job time: 620 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 3, 2002, 08:53:20 ; Search time 46.52 Seconds
(Without alignments)
15.224 Million cell updates/sec

Title: US-09-786-648-2
Perfect score: 39
Sequence: 1 EVPGSQH 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 308740 segs, 101176262 residues

Total number of hits satisfying chosen parameters: 308740

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-Processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending_Patents_AA_New:*
1: /cgn2_6/ptodata/2/paa/PCT_NEW_COMB.pep.*
2: /cgn2_6/ptodata/2/paa/US06_NEW_COMB.pep.*
3: /cgn2_6/ptodata/2/paa/US07_NEW_COMB.pep.*
4: /cgn2_6/ptodata/2/paa/US08_NEW_COMB.pep.*
5: /cgn2_6/ptodata/2/paa/US09_NEW_COMB.pep.*
6: /cgn2_6/ptodata/2/paa/US10_NEW_COMB.pep.*
7: /cgn2_6/ptodata/2/paa/US60_NEW_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	39	100.0	103	6	US-10-110-364-8	Sequence 8, Appl
2	39	100.0	103	6	US-10-110-364-10	Sequence 10, Appl
3	39	100.0	103	6	US-10-110-364-13	Sequence 13, Appl
4	39	100.0	103	6	US-10-110-364-22	Sequence 22, Appl
5	39	100.0	105	6	US-10-110-364-12	Sequence 12, Appl
6	39	100.0	123	6	US-10-110-364-17	Sequence 17, Appl
7	39	100.0	123	6	US-10-110-364-20	Sequence 20, Appl
8	39	100.0	124	6	US-10-110-364-2	Sequence 2, Appl
9	39	100.0	124	6	US-10-110-364-6	Sequence 6, Appl
10	39	100.0	124	6	US-10-110-364-11	Sequence 11, Appl
11	39	100.0	124	6	US-10-110-364-16	Sequence 16, Appl
12	39	100.0	124	6	US-10-110-364-18	Sequence 18, Appl
13	39	100.0	124	6	US-10-110-364-19	Sequence 19, Appl
14	39	100.0	124	6	US-10-110-364-21	Sequence 21, Appl
15	39	100.0	124	6	US-10-110-364-23	Sequence 23, Appl
16	39	100.0	125	6	US-10-110-364-15	Sequence 15, Appl
17	39	100.0	138	6	US-10-141-627-2	Sequence 2, Appl
18	36	92.3	103	6	US-10-110-364-7	Sequence 7, Appl
19	33	84.6	103	6	US-10-110-364-5	Sequence 5, Appl
20	33	84.6	124	6	US-10-110-364-4	Sequence 4, Appl
21	33	84.6	124	6	US-10-110-364-9	Sequence 9, Appl
22	33	84.6	382	5	US-09-809-033A-3	Sequence 3, Appl
23	33	84.6	382	5	US-09-809-033A-4	Sequence 4, Appl
24	33	84.6	412	5	US-09-978-403A-157	Sequence 157, App
25	33	84.6	412	5	US-09-978-544A-157	Sequence 157, App
26	33	84.6	412	5	US-09-978-681A-157	Sequence 157, App

27	33	84.6	412	5	US-09-978-757A-157	Sequence 157, App
28	33	84.6	412	5	US-09-978-564A-157	Sequence 157, App
29	33	84.6	412	5	US-09-999-831A-157	Sequence 157, App
30	33	84.6	412	5	US-09-999-829A-157	Sequence 157, App
31	33	84.6	412	5	US-09-978-375A-157	Sequence 157, App
32	33	84.6	412	5	US-09-978-423A-157	Sequence 157, App
33	33	84.6	412	6	US-10-013-921A-157	Sequence 157, App
34	33	84.6	412	6	US-10-013-929A-157	Sequence 157, App
35	33	84.6	412	6	US-10-013-918A-157	Sequence 157, App
36	33	84.6	412	6	US-10-017-082A-157	Sequence 157, App
37	33	84.6	412	6	US-10-017-085A-157	Sequence 157, App
38	33	84.6	412	6	US-10-013-916A-157	Sequence 157, App
39	33	84.6	412	6	US-10-017-086A-157	Sequence 157, App
40	33	84.6	412	6	US-10-013-925A-157	Sequence 157, App
41	33	84.6	412	6	US-10-017-081A-157	Sequence 157, App
42	33	84.6	412	6	US-10-016-177A-157	Sequence 157, App
43	33	84.6	412	6	US-10-017-084A-157	Sequence 157, App
44	33	84.6	412	6	US-10-013-923A-157	Sequence 157, App
45	33	84.6	473	6	US-10-155-881-26127	Sequence 26127, A

ALIGNMENTS

```

RESULT 1
US-10-110-364-8
; GENERAL INFORMATION:
; APPLICANT: Handley, Harold H.
; APPLICANT: Haaparanta, Tapio
; APPLICANT: Ewall, Karla L.
; TITLE OF INVENTION: ABS TOXIN B SUBUNIT MUTANTS WITH ALTERED
; FILE REFERENCE: ACTBIO.004A
; CURRENT APPLICATION NUMBER: US/10/110.364
; PRIOR FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: PCT/US00/27607
; PRIOR FILING DATE: 2000-10-05
; PRIOR APPLICATION NUMBER: 60/158,561
; PRIOR FILING DATE: 1999-10-08
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 103
; TYPE: PRT
; ORGANISM: Vibrio cholera
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)..(103)
; OTHER INFORMATION: CTB variant from NCBI gene bank GI: 998409
US-10-110-364-8

Query Match 100.0%; Score 39; DB 6; Length 103;
Best Local Similarity 100.0%; Pred. No. 0.79;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0./

OY 1 EVPGSQH 7
Db 51 EVPGSQH 57

RESULT 2
US-10-110-364-10
; GENERAL INFORMATION:
; APPLICANT: Handley, Harold H.
; APPLICANT: Haaparanta, Tapio
; APPLICANT: Ewall, Karla L.
; TITLE OF INVENTION: ABS TOXIN B SUBUNIT MUTANTS WITH ALTERED
; FILE REFERENCE: ACTBIO.004A
; CURRENT APPLICATION NUMBER: US/10/110.364

```

```
; CURRENT FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: PCT/US00/27607
; PRIOR FILING DATE: 2000-10-05
; PRIOR APPLICATION NUMBER: 60/158,561
; PRIOR FILING DATE: 1999-10-08
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 103
; TYPE: PRT
; ORGANISM: Vibrio cholera
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(103)
; OTHER INFORMATION: CTB variant from NCBI gene bank GI: 1421511
US-10-110-364-10
```

```
Query Match          100.0%; Score 39; DB 6; Length 103;
Best Local Similarity 100.0%; Pred. No. 0.79;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 EYPGSOH 7
    |||||
Db  51 EYPGSOH 57
```

```
RESULT 3
US-10-110-364-13
; Sequence 13, Application US/10110364
; GENERAL INFORMATION:
; APPLICANT: Handley, Harold H.
; APPLICANT: Haaparanta, Tapio
; APPLICANT: Ewalt, Karla L.
; TITLE OF INVENTION: ABS TOXIN B SUBUNIT MUTANTS WITH ALTERED
; FILE REFERENCE: ACTBIO.004A
; CURRENT APPLICATION NUMBER: US/10/110.364
; CURRENT FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: PCT/US00/27607
; PRIOR FILING DATE: 2000-10-05
; PRIOR APPLICATION NUMBER: 60/158,561
; PRIOR FILING DATE: 1999-10-08
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 103
; TYPE: PRT
; ORGANISM: Vibrio cholera
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(103)
; OTHER INFORMATION: CTB variant from NCBI gene bank GI: 14215235
; OTHER INFORMATION: (Ogawa 41 R35D).
US-10-110-364-13
```

```
Query Match          100.0%; Score 39; DB 6; Length 103;
Best Local Similarity 100.0%; Pred. No. 0.79;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 EYPGSOH 7
    |||||
Db  51 EYPGSOH 57
```

```
RESULT 4
US-10-110-364-22
; Sequence 22, Application US/10110364
; GENERAL INFORMATION:
; APPLICANT: Handley, Harold H.
; APPLICANT: Haaparanta, Tapio
; APPLICANT: Ewalt, Karla L.
```

```
; TITLE OF INVENTION: ABS TOXIN B SUBUNIT MUTANTS WITH ALTERED
; FILE REFERENCE: ACTBIO.004A
; CURRENT APPLICATION NUMBER: US/10/110.364
; CURRENT FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: PCT/US00/27607
; PRIOR FILING DATE: 2000-10-05
; PRIOR APPLICATION NUMBER: 60/158,561
; PRIOR FILING DATE: 1999-10-08
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 22
; LENGTH: 103
; TYPE: PRT
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(103)
; OTHER INFORMATION: LTB variant from NCBI gene bank GI: 494265.
US-10-110-364-22
```

```
Query Match          100.0%; Score 39; DB 6; Length 103;
Best Local Similarity 100.0%; Pred. No. 0.79;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0,
```

```
OY 1 EYPGSOH 7
    |||||
Db  51 EYPGSOH 57
```

```
RESULT 5
US-10-110-364-12
; Sequence 12, Application US/10110364
; GENERAL INFORMATION:
; APPLICANT: Handley, Harold H.
; APPLICANT: Haaparanta, Tapio
; APPLICANT: Ewalt, Karla L.
; TITLE OF INVENTION: ABS TOXIN B SUBUNIT MUTANTS WITH ALTERED
; FILE REFERENCE: ACTBIO.004A
; CURRENT APPLICATION NUMBER: US/10/110.364
; CURRENT FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: PCT/US00/27607
; PRIOR FILING DATE: 2000-10-05
; PRIOR APPLICATION NUMBER: 60/158,561
; PRIOR FILING DATE: 1999-10-08
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 105
; TYPE: PRT
; ORGANISM: Vibrio cholera
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(105)
; OTHER INFORMATION: CTB variant from NCBI gene bank GI: 2781121 (Ogawa
; OTHER INFORMATION: 41).
US-10-110-364-12
```

```
Query Match          100.0%; Score 39; DB 6; Length 105;
Best Local Similarity 100.0%; Pred. No. 0.81;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0,
```

```
OY 1 EYPGSOH 7
    |||||
Db  52 EYPGSOH 58
```

```
RESULT 6
US-10-110-364-17
; Sequence 17, Application US/10110364
```



```

; GENERAL INFORMATION:
; APPLICANT: Handley, Harold H.
; APPLICANT: Haaparanta, Tapio
; APPLICANT: Ewalt, Karla L.
; TITLE OF INVENTION: ABS TOXIN B SUBUNIT MUTANTS WITH ALTERED
; FILE REFERENCE: ACTBIO.004A
; CURRENT FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: PCT/US00/27607
; PRIOR FILING DATE: 2000-10-05
; PRIOR APPLICATION NUMBER: 60/158,561
; PRIOR FILING DATE: 1999-10-08
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 123
; TYPE: PRT
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(123)
; OTHER INFORMATION: LTB variant from NCBI gene bank GI: 1395122.
US-10-110-364-17

```

```

Query Match          100.0%; Score 39; DB 6; Length 123;
Best Local Similarity 100.0%; Pred. No. 0.96;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 EVPGSOH 7
DB 72 EVPGSOH 78

```

```

RESULT 7
US-10-110-364-20
; Sequence 20, Application US/10110364
; GENERAL INFORMATION:
; APPLICANT: Handley, Harold H.
; APPLICANT: Haaparanta, Tapio
; APPLICANT: Ewalt, Karla L.
; TITLE OF INVENTION: ABS TOXIN B SUBUNIT MUTANTS WITH ALTERED
; FILE REFERENCE: ACTBIO.004A
; CURRENT FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: PCT/US00/27607
; PRIOR FILING DATE: 2000-10-05
; PRIOR APPLICATION NUMBER: 60/158,561
; PRIOR FILING DATE: 1999-10-08
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 20
; LENGTH: 123
; TYPE: PRT
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(123)
; OTHER INFORMATION: LTB variant from NCBI gene bank GI: 223254.
US-10-110-364-20

```

```

Query Match          100.0%; Score 39; DB 6; Length 123;
Best Local Similarity 100.0%; Pred. No. 0.96;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 EVPGSOH 7
DB 71 EVPGSOH 77

```

```

RESULT 8
US-10-110-364-2
; Sequence 2, Application US/10110364
; GENERAL INFORMATION:
; APPLICANT: Handley, Harold H.
; APPLICANT: Haaparanta, Tapio
; APPLICANT: Ewalt, Karla L.
; TITLE OF INVENTION: ABS TOXIN B SUBUNIT MUTANTS WITH ALTERED
; FILE REFERENCE: ACTBIO.004A
; CURRENT FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: PCT/US00/27607
; PRIOR FILING DATE: 2000-10-05
; PRIOR APPLICATION NUMBER: 60/158,561
; PRIOR FILING DATE: 1999-10-08
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 124
; TYPE: PRT
; ORGANISM: Vibrio cholera
US-10-110-364-2

```

```

Query Match          100.0%; Score 39; DB 6; Length 124;
Best Local Similarity 100.0%; Pred. No. 0.96;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 EVPGSOH 7
DB 72 EVPGSOH 78

```

```

RESULT 9
US-10-110-364-6
; Sequence 6, Application US/10110364
; GENERAL INFORMATION:
; APPLICANT: Handley, Harold H.
; APPLICANT: Haaparanta, Tapio
; APPLICANT: Ewalt, Karla L.
; TITLE OF INVENTION: ABS TOXIN B SUBUNIT MUTANTS WITH ALTERED
; FILE REFERENCE: ACTBIO.004A
; CURRENT FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: PCT/US00/27607
; PRIOR FILING DATE: 2000-10-05
; PRIOR APPLICATION NUMBER: 60/158,561
; PRIOR FILING DATE: 1999-10-08
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 124
; TYPE: PRT
; ORGANISM: Vibrio cholera
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(124)
; OTHER INFORMATION: CTB variant from NCBI gene bank GI: 808900.
US-10-110-364-6

```

```

Query Match          100.0%; Score 39; DB 6; Length 124;
Best Local Similarity 100.0%; Pred. No. 0.96;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 EVPGSOH 7
DB 72 EVPGSOH 78

```

```

RESULT 10

```

US-10-110-364-11
; Sequence 11, Application US/10110364
; GENERAL INFORMATION:
; APPLICANT: Handley, Harold H.
; APPLICANT: Haaparanta, Tapio
; APPLICANT: Ewalt, Karla L.
; TITLE OF INVENTION: ABS TOXIN B SUBUNIT MUTANTS WITH ALTERED
; FILE REFERENCE: ACTBIO.004A
; CURRENT APPLICATION NUMBER: US/10/110.364
; PRIOR FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: PCT/US00/27607
; PRIOR FILING DATE: 2000-10-05
; PRIOR APPLICATION NUMBER: 60/158,561
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 124
; TYPE: PRT
; ORGANISM: Vibrio cholera
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(124)
; OTHER INFORMATION: CTB variant from NCBI gene bank GI: 48890 (classic
; US-10-110-364-11

Query Match 100.0%; Score 39; DB 6; Length 124;
Best Local Similarity 100.0%; Pred. No. 0.96;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EYPSGSH 7
1111111
Db 72 EYPSGSH 78

RESULT 11
US-10-110-364-16
; Sequence 16, Application US/10110364
; GENERAL INFORMATION:
; APPLICANT: Handley, Harold H.
; APPLICANT: Haaparanta, Tapio
; APPLICANT: Ewalt, Karla L.
; TITLE OF INVENTION: ABS TOXIN B SUBUNIT MUTANTS WITH ALTERED
; FILE REFERENCE: ACTBIO.004A
; CURRENT APPLICATION NUMBER: US/10/110.364
; PRIOR FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: PCT/US00/27607
; PRIOR FILING DATE: 2000-10-05
; PRIOR APPLICATION NUMBER: 60/158,561
; PRIOR FILING DATE: 1999-10-08
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16
; LENGTH: 124
; TYPE: PRT
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(124)
; OTHER INFORMATION: LTB variant from NCBI gene bank GI: 1169505.
US-10-110-364-16

Query Match 100.0%; Score 39; DB 6; Length 124;
Best Local Similarity 100.0%; Pred. No. 0.96;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EYPSGSH 7
1111111

Db 72 EYPSGSH 78

RESULT 12
US-10-110-364-18
; Sequence 18, Application US/10110364
; GENERAL INFORMATION:
; APPLICANT: Handley, Harold H.
; APPLICANT: Haaparanta, Tapio
; APPLICANT: Ewalt, Karla L.
; TITLE OF INVENTION: ABS TOXIN B SUBUNIT MUTANTS WITH ALTERED
; FILE REFERENCE: ACTBIO.004A
; CURRENT APPLICATION NUMBER: US/10/110.364
; PRIOR FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: PCT/US00/27607
; PRIOR FILING DATE: 2000-10-05
; PRIOR APPLICATION NUMBER: 60/158,561
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 18
; LENGTH: 124
; TYPE: PRT
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(124)
; OTHER INFORMATION: LTB variant from NCBI gene bank GI: 145833.
US-10-110-364-18

Query Match 100.0%; Score 39; DB 6; Length 124;
Best Local Similarity 100.0%; Pred. No. 0.96;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EYPSGSH 7
1111111
Db 72 EYPSGSH 78

RESULT 13
US-10-110-364-19
; Sequence 19, Application US/10110364
; GENERAL INFORMATION:
; APPLICANT: Handley, Harold H.
; APPLICANT: Haaparanta, Tapio
; APPLICANT: Ewalt, Karla L.
; TITLE OF INVENTION: ABS TOXIN B SUBUNIT MUTANTS WITH ALTERED
; FILE REFERENCE: ACTBIO.004A
; CURRENT APPLICATION NUMBER: US/10/110.364
; PRIOR FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: PCT/US00/27607
; PRIOR FILING DATE: 2000-10-05
; PRIOR APPLICATION NUMBER: 60/158,561
; PRIOR FILING DATE: 1999-10-08
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 19
; LENGTH: 124
; TYPE: PRT
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(124)
; OTHER INFORMATION: LTB variant from NCBI gene bank GI: 1648865
US-10-110-364-19

Query Match 100.0%; Score 39; DB 6; Length 124;
Best Local Similarity 100.0%; Pred. No. 0.96;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 EVPGSOH 7
Db 72 EVPGSOH 78

RESULT 14
US-10-110-364-21

; Sequence 21, Application US/10110364
; GENERAL INFORMATION:
; APPLICANT: Handley, Harold H.
; APPLICANT: Haaparanta, Tapio
; APPLICANT: Ewalt, Karla L.
; TITLE OF INVENTION: ABS TOXIN B SUBUNIT MUTANTS WITH ALTERED
; FILE REFERENCE: ACTBIO.004A
; CURRENT APPLICATION NUMBER: US/10/110,364
; CURRENT FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: PCT/US00/27607
; PRIOR FILING DATE: 2000-10-05
; PRIOR APPLICATION NUMBER: 60/158,561
; PRIOR FILING DATE: 1999-10-08
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 21
; LENGTH: 124
; TYPE: PRT
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(124)
; OTHER INFORMATION: LTB variant from NCBI gene bank GI: 408996.
US-10-110-364-21

Query Match 100.0%; Score 39; DB 6; Length 124;
Best Local Similarity 100.0%; Pred. No. 0.96;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EVPGSOH 7
Db 72 EVPGSOH 78

RESULT 15
US-10-110-364-23

; Sequence 23, Application US/10110364
; GENERAL INFORMATION:
; APPLICANT: Handley, Harold H.
; APPLICANT: Haaparanta, Tapio
; APPLICANT: Ewalt, Karla L.
; TITLE OF INVENTION: ABS TOXIN B SUBUNIT MUTANTS WITH ALTERED
; FILE REFERENCE: ACTBIO.004A
; CURRENT APPLICATION NUMBER: US/10/110,364
; CURRENT FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: PCT/US00/27607
; PRIOR FILING DATE: 2000-10-05
; PRIOR APPLICATION NUMBER: 60/158,561
; PRIOR FILING DATE: 1999-10-08
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 23
; LENGTH: 124
; TYPE: PRT
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(124)
; OTHER INFORMATION: LTB variant from NCBI gene bank GI: 69630.
US-10-110-364-23

Query Match 100.0%; Score 39; DB 6; Length 124;
Best Local Similarity 100.0%; Pred. No. 0.96;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EVPGSOH 7
Db 72 EVPGSOH 78

Search completed: July 3, 2002, 09:04:12
Job time: 652 sec

Mon Jul 8 07:51:43 2002

us-09-786-648-2.rapn

Page 6

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 3, 2002, 08:52:40 ; Search time 46.57 Seconds
(without alignments)
14.443 Million cell updates/sec

Title: US-09-786-648-2

Perfect score: 39

Sequence: 1 EYPGSGH 7

Scoring table: BIOSUM62
Gapop 10.0, Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 08
Maximum Match 1008
Listing first 45 summaries

Database :

1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	39	100.0	124	1	XVCB
2	39	100.0	124	1	OLECB
3	33	84.6	272	2	T14755
4	33	84.6	367	2	T01751
5	32	82.1	137	2	G96017
6	32	82.1	455	1	A69753
7	32	82.1	472	2	AG0432
8	32	82.1	534	2	S44886
9	31	79.5	137	2	B27586
10	31	79.5	242	2	T34767
11	31	79.5	260	2	C96827
12	31	79.5	273	2	B48820
13	31	79.5	274	2	G84353
14	31	79.5	353	2	H75446
15	31	79.5	374	2	D81715
16	31	79.5	414	2	A37133
17	31	79.5	414	2	I48975
18	31	79.5	432	2	T46725
19	31	79.5	432	2	T51020
20	31	79.5	459	2	D34791
21	31	79.5	489	2	T26069
22	31	79.5	703	2	AG0242
23	31	79.5	708	2	AG2315
24	31	79.5	755	2	D95842
25	31	79.5	1028	2	A96719
26	31	79.5	1228	2	C98219
27	31	79.5	1228	2	AG3067
28	31	79.5	1258	2	T29041
29	31	79.5	1259	2	A43425

30	30	76.9	148	2	B72782	hypothetical prote
31	30	76.9	153	2	T31701	hypothetical prote
32	30	76.9	182	2	D83638	conserved hypotnet
33	30	76.9	262	2	C34791	interleukin-7 rece
34	30	76.9	270	2	D83072	conserved hypotnet
35	30	76.9	291	2	A12241	hypothetical prote
36	30	76.9	298	2	B34791	interleukin-7 rece
37	30	76.9	304	2	F84169	hypothetical prote
38	30	76.9	336	2	A72247	DNA-directed RNA p
39	30	76.9	373	2	T47115	probable 4'-carboxy
40	30	76.9	375	2	T35015	probable 3'-oxadip
41	30	76.9	376	1	S17246	choistamate synthas
42	30	76.9	412	2	T47142	hypothetical prote
43	30	76.9	431	2	A89761	hypothetical prote
44	30	76.9	436	2	B70321	flavocytochrome C
45	30	76.9	440	2	T44138	hypothetical prote

ALIGNMENTS

RESULT 1
XVCB
Cholera enterotoxin, B chain precursor VC1456 [validated] - Vibrio cholerae (strain N
N: Alternate names: enterotoxin beta chain
C: Species: Vibrio cholerae
C: Date: 24-Apr-1984 #sequence, revision 01-Sep-2000 #text, change 02-Feb-2001
C: Accession: S14624; S39238; S39241; H82196; JCI078; S17666; PC1010; A05130; A01819,
R: Dams, E.; de Wolf, M.; Dierick, W.
submitted to the EMBL Data Library, March 1991
A: Description: Correction of the cholera toxin nucleotide sequence of the Vibrio chol
A: Reference number: S14623
A: Accession: S14624
A: Molecule type: DNA
A: Residues: 1-124 <DM>
A: Cross-references: EMBL:X58786; NID:G48420; PIDN:CAA1593.1; PID:G48422
A: Experimental source: strain 2125
R: Lebens, M.; Holmgren, J.
submitted to the EMBL Data Library, November 1993
A: Description: Structure and arrangement of the Cholera toxin genes in vibrio Cholera
A: Reference number: S39238
A: Accession: S39238
A: Molecule type: DNA
A: Residues: 1-124 <LEB>
A: Cross-references: EMBL:X76390; NID:G433856; PIDN:CAA53973.1; PID:G433857
A: Accession: S39241
A: Molecule type: DNA
A: Residues: 1-124 <LEM>
A: Cross-references: EMBL:X76391; NID:G433859; PIDN:CAA53976.1; PID:G433861
R: Heideberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwin, M.L.; Dodson,
Chardson, D.; Ermoleva, M.D.; Vamathevan, J.; Baas, S.; Qin, H.; Dragol, I.; Sellc
1, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A: Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae
A: Reference number: A82035; MID:20406833
A: Accession: H82196
A: Status: preliminary
A: Molecule type: DNA
A: Residues: 1-124 <HEI>
A: Cross-references: GB:AE004224; GB:AE003852; NID:G9655952; PIDN:AAF94613.1; GSP
R: Shi, C.H.; Cao, C.; Zhang, J.S.; Ma, Q.J.
Chinese Nucleotide sequence analysis of the gene encoding the classical biotype
A: Title: Nucleotide sequence analysis of the gene encoding the classical biotype
A: Reference number: JCI078
A: Accession: JCI078
A: Molecule type: DNA
A: Residues: 1-20, 'Q', 22-31, 'Q', 33-38, 'H', 40-49, 'G', 51-67, 'T', 69-124 <SHI>
A: Experimental source: Classical biotype strain 569B
R: Dams, E.; de Wolf, M.; Dierick, W.
Biochim. Biophys. Acta 1090, 139-141, 1991
A: Title: Nucleotide sequence analysis of the CT operon of the Vibrio cholerae
A: Reference number: S17665; MID:91355224

A:Accession: S17666
A:Molecule type: DNA
A:Residues: 1-38, 'H', 40-67, 'T', 69-124 <DA2>
A:Cross-references: EMBL:X58785; NID:948888; PIDN:CAA41591.1; PID:948890
R:Ma, O.J.; Liu, C.X.; Xiong, L.S.; Yu, X.Q.
Sci. Sin. B Chem. Biol. Agric. Med. Earth Sci. 34, 274-280, 1991
A:Title: B subunit of cholera toxin produced in *Escherichia coli*.
A:Reference number: PC1010
A:Accession: PC1010
A:Molecule type: Protein
A:Residues: 22-38, 'H', 40-41 <MAO>
R:Meekalanos, J.J.; Swartz, D.J.; Pearson, G.D.N.; Harford, N.; Groyne, F.; de Wilde, M.
Nature 306, 551-557, 1983
A:Accession: A05130
A:Reference number: A05130
A:Molecule type: DNA
A:Residues: 1-32, 'S', 34-74, 'S', 76-124 <MEK>
A:Cross-references: GB:X00171; NID:948347; PIDN:CAA24996.1; PID:9758351
R:Kurosaki, A.; Martel, D.E.; Peterson, J.W.
J. Biol. Chem. 252, 7257-7264, 1977
A:Title: Covalent structure of the beta chain of cholera enterotoxin.
A:Reference number: A01819; MUID:78005537
A:Accession: A01819
A:Molecule type: Protein
A:Residues: 22-38, 'H', 40-42, 'N', 44-67, 'T', 69-90, 'N', 92-124 <KUR>
R:Lat, C.Y.
J. Biol. Chem. 252, 7249-7256, 1977
A:Title: Determination of the primary structure of cholera toxin B subunit.
A:Reference number: A38033; MUID:78005536
A:Accession: A38033
A:Molecule type: Protein
A:Residues: 22-38, 'H', 40-42, 'N', 44-67, 'T', 69, 'E', 71-90, 'N', 92-124 <LAI>
A:Note: The difference at residue 70 may be due to deamidation during preparation
R:Nakashima, Y.; Napolitkowski, P.; Schafer, D.E.; Konigsberg, W.H.
FEBS Lett. 68, 275-278, 1976
A:Title: Primary structure of the B subunit of cholera enterotoxin.
A:Reference number: A38034; MUID:77026365
A:Accession: A38034
A:Molecule type: Protein
A:Residues: 22-38, 'H', 40-67, 'T', 69, 'E', 71, 'QS', 74-75, 'VE', 78-86, 'Q', 88-99, 'Q', 101-103, 'Q'
R:Takao, T.; Melanabe, H.; Shimonishi, Y.
Eur. J. Biochem. 146, 503-508, 1985
A:Title: Facile identification of protein sequences by mass spectrometry.
A:Reference number: A21910; MUID:85126976
A:Accession: A21910
A:Molecule type: Protein
A:Residues: 22-38, 'H', 40-42, 'N', 44-67, 'T', 69-90, 'N', 92-124 <TAK>
A:Experimental source: biotype Inaba 569B
A:Note: Asn-65 was partially deaminated to Asp
C:Comment: The authors translated the codon TCA for residue 33 as Tyr.
C:Genetics:
A:Gene: VC1456
A:Map position: 1
C:Complex: the cholera enterotoxin molecule contains three kinds of chains; an alpha and
beta noncovalently with the subunit B, an aggregate of five beta chains
C:Function:
A:Description: involved in binding of the toxin to cell membranes
C:Superfamily: cholera enterotoxin beta chain
C:Keywords: enterotoxin; toxin
F:1-21/Domain: signal sequence #status predicted <SIG>
F:22-124/Product: cholera enterotoxin chain B #status predicted <MAT>
F:30-107/Disulfide bonds: #status experimental

Query Match 100.0%; Score 39; DB 1; Length 124;
Best Local Similarity 100.0%; Pred. No. 0.5;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EVPGSOH 7
DB 72 EVPGSOH 78

RESULT 2
OLECB
heat-labile enterotoxin chain B precursor - *Escherichia coli*
C:Species: *Escherichia coli*
C:Date: 29-Jun-1981 #sequence revision 29-Jun-1981 #text change 18-Jun-1999
C:Accession: A01820; B26946; I41194; I41287; I67644; A61475
R:Dallas, W.S.; Falkow, S.
Nature 288, 499-501, 1980
A:Title: Amino acid sequence homology between cholera toxin and *Escherichia coli*
A:Reference number: A01820; MUID:81074965
A:Accession: A01820
A:Molecule type: mRNA
A:Residues: 1-124 <DAL>
R:Yamamoto, T.; Gojobori, T.; Yokota, T.
J. Bacteriol. 169, 1352-1357, 1987
A:Title: Evolutionary origin of pathogenic determinants in enterotoxigenic *Escherichia coli*
A:Reference number: A26946; MUID:87137303
A:Accession: B26946
A:Molecule type: DNA
A:Residues: 1-27, 'E', 29-63, 'K', 65-124 <YAM>
A:Cross-references: EMBL:M15363; NID:948335; PIDN:AAA24792.1; PID:948336
R:Leong, J.; Vinal, A.C.; Dallas, W.S.
Infect. Immun. 48, 73-77, 1985
A:Title: Nucleotide sequence comparison between heat-labile toxin B-subunit cistrons
A:Reference number: I41194; MUID:85156481
A:Accession: I41194
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-5, 'F', 7-17, 'C', 19-24, 'S', 26-27, 'E', 29-33, 'H', 35-63, 'K', 65-66, 'A', 68-122
A:Cross-references: GB:M17874; NID:9445830; PIDN:AAA98064.1; PID:9445831
R:Idrahimi, I.; Gentz, R.
J. Biol. Chem. 262, 10189-10194, 1987
A:Title: A functional interaction between the signal peptide and the translation apparatus
A:Reference number: I41287; MUID:87280041
A:Accession: I41287
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-22 <RE2>
A:Cross-references: GB:M17101; NID:9446375; PIDN:AAA23973.1; PID:9446376
R:Imone, T.; Tsuji, T.; Koto, M.; Imamura, S.; Miyama, A.
FEBS Microbiol. Lett. 108, 157-161, 1993
A:Title: Amino acid sequence of heat-labile enterotoxin from chicken enterotoxigenic
A:Reference number: I5342; MUID:93252225
A:Accession: I67644
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-17, 'C', 19, 'Y', 21-24, 'S', 26-27, 'E', 29-63, 'K', 65-66, 'A', 68-122, 'E', 124 <R>
A:Cross-references: GB:S60731; NID:9408994; PIDN:AA60441.1; PID:9408996
R:Tsuji, T.; Lida, T.; Honda, T.; Mawatani, T.; Nagahama, M.; Sakurai, J.; Wada, K.
Microb. Pathog. 2, 381-390, 1987
A:Title: A unique amino acid sequence of the B subunit of a heat-labile enterotoxin 1
A:Reference number: A61475; MUID:89180953
A:Accession: A61475
A:Molecule type: Protein
A:Residues: 22-24, 'S', 26-27, 'E', 29-63, 'K', 65-66, 'A', 68-95, 'A', 97-122, 'E', 124 <TSU>
A:Experimental source: strain 240-3
C:Complex: the heat-labile enterotoxin molecule contains one A chain and five or six
C:Function:
A:Description: the biological activity of the toxin is produced by the A chain, which
C:Superfamily: cholera enterotoxin beta chain
C:Keywords: enterotoxin
F:1-21/Domain: signal sequence #status predicted <SIG>
F:22-124/Product: heat-labile enterotoxin chain B #status predicted <MAT>
F:30-107/Disulfide bonds: #status predicted

Query Match 100.0%; Score 39; DB 1; Length 124;
Best Local Similarity 100.0%; Pred. No. 0.5;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EVPGSOH 7

Db 72 EVPGSQH 78
|||||

RESULT 3

T14755

hypothetical protein DKFZP564A0122.1 - human

C:Species: Homo sapiens (man)
C:Date: 20-Sep-1999 #sequence.revision 20-Sep-1999 #text.change 20-Sep-1999

C:Accession: T14755

R:Wambutt, R.; Heubner, D.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, August 1999

A:Reference number: 218181

A:Accession: T14755

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-272 <NAM>

A:Cross-references: EMBL:AL10209

A:Experimental source: fetal brain; clone DKFZP564A0122

A:Note: DKFZP564A0122.1

Query Match 84.6%; Score 33; DB 2; Length 272;
Best Local Similarity 71.4%; Pred. NO. 22;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 EVPGSQH 7
|||||
Db 246 EDPGSEH 252

RESULT 4

T01751

giberellin 20-oxidase - common tobacco

N:Alternate names: Ntc16 protein

C:Species: Nicotiana tabacum (common tobacco)

C:Date: 19-Feb-1999 #sequence.revision 19-Feb-1999 #text.change 20-Jun-2000

C:Accession: T01751

R:Tanaka-Ieguchi, M.; Itoh, H.; Oyama, N.; Koshioka, M.; Matsuo, M.
submitted to the EMBL Data Library, July 1998

A:Description: Over-expression of a tobacco homeobox gene, NTH15, decreases the expres

A:Reference number: 214418

A:Accession: T01751

A:Status: preliminary; translated from GB/EMBL/DDB

A:Molecule type: mRNA

A:Residues: 1-367 <TAN>

A:Cross-references: EMBL:AB016084

A:Genetics:

A:Gene: NTC16

C:Superfamily: 1-aminocyclopropane-1-carboxylate oxidase

Query Match 84.6%; Score 33; DB 2; Length 367;
Best Local Similarity 85.7%; Pred. NO. 30;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 EVPGSQH 7
|||||
Db 149 EPPSSQH 155

RESULT 5

G96017

conserved hypothetical exported protein SMD20700 [imported] - Sinorhizobium meliloti (st

C:Species: Sinorhizobium meliloti

C:Date: 24-Aug-2001 #sequence.revision 24-Aug-2001 #text.change 30-Sep-2001

C:Accession: G96017

R:Finan, T.M.; Weidner, S.; Wong, K.; Buhrmester, J.; Chain, P.; Vorholter, F.J.; Hernan

Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001

A:Title: The complete sequence of the 1,683-kb PSYMB megaplasmid from the N2-fixing endo

A:Reference number: A95842; MUID:21396508; PMID:11481431

A:Accession: G96017

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-137 <KUR>

A:Cross-references: GB:AL591985; PIDN:CAC49807.1; PID:q15141295; GSPDB:GN00167

A:Experimental source: strain 1021, megaplasmid PSYMB

R:Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abolte, P.; Ampe, F.; Barloy-Hub

pel, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.

L.; Hyman, R.W.; Jones, T.

Science 293, 668-672, 2001

A:Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelu

heault, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh,

A:Title: The composite genome of the legume symbiont Sinorhizobium meliloti.

A:Reference number: A96039; MUID:21368234; PMID:11474104

A:Contents: annotation

C:Genetics:

A:Gene: SMD20700

A:Genome: plasmid

Query Match 82.1%; Score 32; DB 2; Length 137;
Best Local Similarity 85.7%; Pred. NO. 17;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 EVPGSQH 7
|||||
Db 98 EDPGSEH 104

RESULT 6

A69753

glucarate dehydratase (EC 4.2.1.40) - Bacillus subtilis

C:Species: Bacillus subtilis

C:Date: 10-Sep-1999 #sequence.revision 10-Sep-1999 #text.change 16-Jun-2000

C:Accession: A69753

R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Ber

C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.;

A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari,

Nature 390, 249-256, 1997

A:Authors: Foulger, D.; Fritze, C.; Fujita, M.; Fuma, S.; Galliz, A.; Gal

lech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holtsappel, S.; Hosono, S.; Hullo, M

Koetter, P.; Konigstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardino

A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Mau

y, M.; Ogawa, K.; Ogata, A.; Oudega, B.; Park, S.H.; Paro, V.; Pohl, T.M.; Portee

Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadale, Y.; Sato, T.; Scanl

A:Authors: Schleich, S.; Schroeter, R.; Scottone, F.; Sekiguchi, J.; Sekowska, A.; Se

akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiya

T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida

A:Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A.

A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis

A:Reference number: A69580; MUID:98044033

A:Accession: A69753

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-455 <KUN>

A:Cross-references: GB:299105; GB:AL009126; NID:q2632457; PIDN:CAB12043.1; PID:q26325

A:Experimental source: strain 168

C:Genetics:

A:Gene: ycbF

C:Superfamily: glucarate dehydratase

C:Keywords: carbon-oxygen lyase; hydro-lyase

Query Match 82.1%; Score 32; DB 1; Length 455;
Best Local Similarity 71.4%; Pred. NO. 62;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 EVPGSQH 7
|||||
Db 64 EDPGSEH 70

RESULT 7

AG0432

glutamate synthase (NADPH) (EC 1.4.1.13) small chain [imported] - *Yersinia pestis* (strain C:Species: *Yersinia pestis*)
 C:Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 27-Nov-2001
 C:Accession: AC0432
 R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Tlball, R.W.; Holden, M.T.G.; Prentice, M.B.; deno-Tarrag, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.; 11, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrett, Nature 413, 523-527, 2001
 A:Title: Genome sequence of *Yersinia pestis*, the causative agent of plague.
 A:Reference number: AB0001; MUID:21470413; PMID:11586360
 A:Accession: AC0432
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-472 <KUR>
 A:Cross-references: GB:AL590842; PIDN:CAC92787.1; PID:q15981480; GSPDB:GN00175
 C:Genetics:
 A:Gene: q1td
 C:Superfamily: glutamate synthase small chain
 C:Keywords: oxidoreductase

Query Match 82.1% Score 32; DB 2; Length 472;
 Best Local Similarity 57.1% Pred. No. 65;
 Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 1 EYVGSQH 7
 ::|||:|
 Db 382 QIPGSEH 388

RESULT 8
 S4486
 ZK112.1 protein - *Caenorhabditis elegans*
 C:Species: *Caenorhabditis elegans*
 C:Date: 14-Sep-1994 #sequence_revision 12-May-1995 #text_change 23-Mar-2001
 C:Accession: S4486
 R:Du, Z.
 submitted to the EMBL Data Library, May 1993
 A:Description: Sequence of the C. elegans cosmid ZK112.
 A:Reference number: S4486
 A:Accession: S4486
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-534 <DUZ>
 A:Cross-references: EMBL:L14324; NID:q289740; PID:q289741
 C:Genetics:
 A:Introns: 25/3; 65/2; 196/2; 249/1; 275/1; 385/2; 415/2
 C:Superfamily: *Caenorhabditis elegans* ZK688.6 protein

Query Match 82.1% Score 32; DB 2; Length 534;
 Best Local Similarity 71.4% Pred. No. 74;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 EYVGSQH 7
 ||:||||
 Db 442 EIPGSAH 448

RESULT 9
 B27586
 hypothetical protein - *Mycobacterium leprae*
 C:Species: *Mycobacterium leprae*
 C:Date: 08-Mar-1989 #sequence_revision 08-Mar-1989 #text_change 26-May-2000
 C:Accession: B27586
 R:Booth, R.J.; Harris, D.P.; Love, J.M.; Watson, J.D.
 J. Immunol. 140, 597-601, 1988
 A:Title: Antigenic proteins of *Mycobacterium leprae*. Complete sequence of the gene for
 A:Reference number: A92821; MUID:88088878
 A:Accession: B27586
 A:Molecule type: DNA
 A:Residues: 1-137 <BOO>
 A:Cross-references: GB:M19058; NID:q149919; PIDN:AAA88230.1; PID:q1196505

C:Superfamily: *Mycobacterium leprae* hypothetical 15.2K protein

Query Match 79.5% Score 31; DB 2; Length 137;
 Best Local Similarity 83.3% Pred. No. 28;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 VPGSQH 7
 |||:|
 Db 127 VPGNOH 132

RESULT 10
 T34767
 hypothetical protein SC2A11.21c SC2A11.21c - *Streptomyces coelicolor*
 C:Species: *Streptomyces coelicolor*
 C:Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 05-Nov-1999
 C:Accession: T34767
 R:Murphy, L.; Harris, D.; Parkhill, J.; Barrett, B.G.; Rajandream, M.A.
 submitted to the EMBL Data Library, August 1998
 A:Reference number: T34767
 A:Accession: T34767
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-242 <MUR>
 A:Cross-references: EMBL:AL031184; PIDN:CAA20190.1; GSPDB:GN00070; SCOEDB:SC2A11.21c
 A:Experimental source: strain A3(2)
 C:Genetics:
 A:Gene: SCOEDB:SC2A11.21c

Query Match 79.5% Score 31; DB 2; Length 242;
 Best Local Similarity 71.4% Pred. No. 52;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 EYVGSQH 7
 ||||:|
 Db 205 EYVGDH 211

RESULT 11
 C96827
 protein F20B17.2 [imported] - *Arabidopsis thaliana*
 C:Species: *Arabidopsis thaliana* (mouse-ear cress)
 C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
 C:Accession: C96827
 R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alon
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar,
 ansen, N.F.; Hughes, B.; Huizar, L.
 Nature 408, 816-820, 2000
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim,
 C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Malt, R.; Marzla
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A:Authors: Salberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallo
 ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A:Title: Sequence and analysis of chromosome 1 of the plant *Arabidopsis*.
 A:Reference number: A66141; MUID:21016719
 A:Accession: C96827
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-260 <STO>
 A:Cross-references: GB:AE005173; NID:q7715588; PIDN:AAF68106.1; GSPDB:GN00141
 C:Genetics:
 A:Gene: F20B17.2
 A:Map position: 1

Query Match 79.5% Score 31; DB 2; Length 260;
 Best Local Similarity 71.4% Pred. No. 56;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 EYVGSQH 7
 :|||:|

DB 66 KVPKOH 72

RESULT 12

B48820

homeobox protein (clone NvHBox-4) - eastern newt

C:Species: *Notophthalmus viridescens*, *Triturus viridescens* (eastern newt)

C:Date: 01-Dec-1993 #sequence_revision 18-Nov-1994 #text_change 24-Sep-1999

C:Accession: B48820

R:Beauchemin, M.; Savard, P.

Dev. Biol. 154, 55-65, 1992

A:Title: Two distal-less related homeobox-containing genes expressed in regeneration bio

A:Reference number: A48820; MUID:93050784

A:Accession: B48820

A:Status: preliminary

A:Molecule type: nucleic acid

A:Residues: 1-273 <BEA>

A:Cross-references: GB:X63531; GB:SA7223; NID:9432377; PIDN:CAA45094.1; PID:9432378

A:Note: sequence extracted from NCBI backbone (NCBIN:117052, NCBIP:117053)

C:Superfamily: unassigned homeobox proteins; homeobox homology

C:Keywords: DNA binding; homeobox; nucleus; transcription regulation

F:126-182/Domain: homeobox homology <HOX>

Query Match

Best Local Similarity 79.5%; Score 31; DB 2; Length 273;

Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 EYPGSOH 7

DB 188 EYPGMEH 194

|||||

RESULT 13

G84353

hypothetical protein Vng2034h [imported] - *Halobacterium* sp. NRC-1C:Species: *Halobacterium* sp. NRC-1

C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001

C:Accession: G84353

R:Ng, W.V.; Kennedy, S.P.; Mahaitas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky, S.

; Leithausner, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jabid

Jung, K.H.; Alm, S.; Freitas, T.

Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000

A:Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; Li

A:Title: Genome sequence of *Halobacterium* species NRC-1.

A:Reference number: A84160; MUID:20504483

A:Accession: G84353

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-274 <STO>

A:Cross-references: GB:AE004437; NID:910581460; PIDN:AMG20195.1; GSPDB:GMO0138

C:Genetics:

A:Gene: VNG2034H

Query Match

Best Local Similarity 79.5%; Score 31; DB 2; Length 274;

Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 EYPGSOH 7

DB 117 EAPGDH 123

|||||

RESULT 14

H75446

(S)-2-hydroxy-acid oxidase - *Deinococcus radiodurans* (strain R1)C:Species: *Deinococcus radiodurans*

C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 28-Jul-2000

C:Accession: H75446

R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;

M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma

S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.

Science 286, 1571-1577, 1999

A:Title: Genome sequence of the radioresistant bacterium *Deinococcus radiodurans* R1.

A:Reference number: A75250; MUID:20036896

A:Accession: H75446

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-353 <MHI>

A:Cross-references: GB:AE001954; GB:AE000513; NID:96458751; PIDN:AAF10604.1; PID:9645

A:Experimental source: strain R1

C:Genetics:

A:Gene: DR1031

A:Map position: 1

C:Superfamily: (S)-2-hydroxy-acid oxidase; (S)-2-hydroxy-acid oxidase homology

F:3-297/Domain: (S)-2-hydroxy-acid oxidase homology <2HT>

Query Match

Best Local Similarity 79.5%; Score 31; DB 2; Length 353;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 VPGSOH 7

DB 189 VPGSEH 194

|||||

RESULT 15

D81715

conserved hypothetical protein TC0328 [imported] - *Chlamydia muridarum* (strain Nigg)C:Species: *Chlamydia muridarum*, *Chlamydia trachomatis* Morn

C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 11-May-2000

C:Accession: D81715

R:Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hicke

, C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzbe

Nucleic Acids Res. 28, 1397-1406, 2000

A:Title: Genome sequences of *Chlamydia trachomatis* Morn and *Chlamydia pneumoniae* AR39

A:Reference number: A81500; MUID:20150255

A:Accession: D81715

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-374 <TEY>

A:Cross-references: GB:AE002300; GB:AE002160; NID:97190365; PIDN:AAF39192.1; PID:9719

C:Genetics:

A:Experimental source: strain Nigg (Morn)

A:Gene: TC0328

Query Match

Best Local Similarity 79.5%; Score 31; DB 2; Length 374;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 VPGSOH 7

DB 327 LPGSOH 332

|||||

Search completed: July 3, 2002, 08:56:32
Job time: 232 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 3, 2002, 08:55:00 ; Search time 21.51 Seconds
(Without alignments)
12.601 Million cell updates/sec

Title: US-09-786-648-2
Perfect score: 39
Sequence: 1 EYFGSQH 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	39	100.0	124	1	CHTB_VIRCH
2	39	100.0	124	1	ELBH_ECOLI
3	39	100.0	124	1	ELBP_ECOLI
4	32	82.1	455	1	GUDH_BACSU
5	32	82.1	534	1	YOG1_CAEEL
6	32	82.1	666	1	PD14_MOUSE
7	32	82.1	765	1	SIIM_MOUSE
8	31	79.5	137	1	YH96_MYCLE
9	31	79.5	219	1	CITDA_HUMAN
10	31	79.5	240	1	YB5A_THEMA
11	31	79.5	273	1	DLX3_NORVI
12	31	79.5	274	1	DLX3_PLEMA
13	31	79.5	280	1	DLX3_AMBME
14	31	79.5	414	1	COT2_BOVIN
15	31	79.5	414	1	COT2_HUMAN
16	31	79.5	414	1	COT2_MOUSE
17	31	79.5	414	1	COT2_RAT
18	31	79.5	432	1	AROC_NEUCR
19	31	79.5	459	1	IL7R_MOUSE
20	31	79.5	710	1	IRAL_MOUSE
21	31	79.5	743	1	BGAL_THRET
22	30	76.9	182	1	KPTA_PSEAE
23	30	76.9	336	1	RPOA_THEMA
24	30	76.9	376	1	AROC_YEAST
25	30	76.9	459	1	IL7R_HUMAN
26	30	76.9	500	1	GABT_HUMAN
27	30	76.9	504	1	A37C_DROME
28	30	76.9	573	1	AMH2_HUMAN
29	30	76.9	616	1	REFX5_HUMAN
30	30	76.9	622	1	SR68_CAEEL
31	30	76.9	766	1	SIIM_HUMAN
32	30	76.9	775	1	THIL_SCHPO
33	30	76.9	805	1	AHR_MOUSE

34	30	76.9	808	1	PLD_TOBAC
35	30	76.9	853	1	AHR_RAT
36	30	76.9	954	1	ST14_HUMAN
37	30	76.9	954	1	DRP2_HUMAN
38	30	76.9	1286	1	PATC_DROME
39	30	76.9	1538	1	LHR_ECOLI
40	30	76.9	1638	1	BRM_DROME
41	30	76.9	1808	1	TENA_CHICK
42	29	74.4	277	1	ZDKG_CONSP
43	29	74.4	395	1	INX3_DROME
44	29	74.4	404	1	Y4XM_RHISN
45	29	74.4	586	1	RRPO_BWYVF

ALIGNMENTS

RESULT 1	ID	CHTB_VIRCH	STANDARD	PRT	124 AA.
AC	P01556	Q9JQ02			
DT	21-JUL-1986	(Rel. 01, Created)			
DT	13-AUG-1987	(Rel. 05, Last sequence update)			
DT	16-OCT-2001	(Rel. 40, Last annotation update)			
DE	Cholera enterotoxin, beta chain precursor.				
GN	CITX OR TOXB OR VC1456.				
OC	Vibrio cholerae.				
OS	Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.				
OX	NCBI_TaxID=666;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=84061784; PubMed=6315707;				
RA	Lockman H., Kaper J.B.;				
RT	"Nucleotide sequence analysis of the A2 and B subunits of Vibrio				
RT	cholerae enterotoxin.";				
RL	J. Biol. Chem. 258:13722-13726(1983).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=EL TOR 2125;				
RX	MEDLINE=84068199; PubMed=6646234;				
RA	Mekalanos J.J., Swartz D.J., Pearson G.D.N., Harford N., Groyne F.,				
RA	de Wilde M.;				
RT	"Cholera toxin genes: nucleotide sequence, deletion analysis and				
RT	vaccine development.";				
RL	Nature 306:551-557(1983).				
RN	[3]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=EL TOR 2125;				
RX	MEDLINE=94237453; PubMed=8181723;				
RA	Lebens M., Holmgren J.;				
RT	"Structure and arrangement of the cholera toxin genes in Vibrio				
RT	cholerae O139.";				
RL	FEMS Microbiol. Lett. 117:197-202(1994).				
RN	[5]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=1854 / O139-BENGAL;				
RA	Yamamoto K., Do V.G.R.F., Xu M., Iida T., Miwatani T., Albert M.J.,				
RA	Honda T.;				
RL	submitted (May-1994) to the EMBL/GenBank/DBJ databases.				
RN	[6]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=EL TOR N16961 / SEROTYPE O1;				
RX	MEDLINE=20406833; PubMed=10952301;				
RA	Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwinn M.L.,				
RA	Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,				
RA	Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,				
RA	Ernolleva M.D., Vamathevan J., Bass S., Qin H., Dragoti I., Sellers P.,				
RA	McDonald L., Uitterback T., Fleischmann R.D., Nierman W.C., White O.,				

RA Salzborg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
 RA Fraser C.M.;
 RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio
 RT cholerae.";
 RL Nature 406:477-483(2000).
 RN [7]
 RP SEQUENCE OF 22-124.
 RX MEDLINE=78005537; PubMed=903363;
 RA Kurosky A., Markel D.E., Peterson J.W.;
 RT "Covalent structure of the beta chain of cholera enterotoxin.";
 RL J. Biol. Chem. 252:7257-7264(1977).
 RN [8]
 RP SEQUENCE OF 22-124.
 RX MEDLINE=78005536; PubMed=903362;
 RA Lai C.-Y.;
 RT "Determination of the primary structure of cholera toxin B subunit.";
 RL J. Biol. Chem. 252:7249-7256(1977).
 RN [9]
 RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS).
 RX MEDLINE=94272319; PubMed=8003954;
 RA Merritt E.A., Sarfaty S., van den Akker F., L'Hoir C., Martini J.A.,
 RT "Crystal structure of cholera toxin B-pentamer bound to receptor GM1
 RT pentasaccharide.";
 RL Protein Sci. 3:166-175(1994).
 RN [10]
 RP X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS).
 RX MEDLINE=95387394; PubMed=7658472;
 RA Zhang R.G., Westbrook M.L., Westbrook E.M., Scott D.L., Oclowski Z.,
 RT "The 2.4 A crystal structure of cholera toxin B subunit pentamer:
 RT choleragenoid.";
 RL J. Mol. Biol. 251:550-562(1995).
 RN [11]
 RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).
 RC STRAIN-OGAMA 41 / CLASSICAL BIOTYPE;
 RX MEDLINE=97376625; PubMed=9232653;
 RA Merritt E.A., Sarfaty S., Jobling M.G., Chang T., Holmes R.K.,
 RT "Structural studies of receptor binding by cholera toxin mutants.";
 RL Protein Sci. 6:1516-1528(1997).
 CC -1- FUNCTION: THE BETA CHAIN AGGREGATE (B SUBUNIT) IS INVOLVED IN
 CC BINDING TO CELL MEMBRANES.
 CC -1- SUBUNIT: CONTAINS 3 KINDS OF CHAINS. AN ALPHA AND A GAMMA CHAIN
 CC (FROM THE SAME PRECURSOR MOLECULE), LINKED BY AN INTERCHAIN
 CC DISULFIDE BOND, ASSOCIATE NONCOVALENTLY WITH AN AGGREGATE OF 4 TO
 CC 6 BETA CHAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: X00171; CAA24996.1; -
 DR EMBL: K01170; AAA27573.1; -
 DR EMBL: D30053; BAA06291.1; -
 DR EMBL: X58786; CAA41593.1; -
 DR EMBL: X76390; CAA53973.1; -
 DR EMBL: X76391; CAA53976.1; -
 DR EMBL: AE004224; AAF94613.1; -
 DR PIR: A01819; XVCB.
 DR PIR: A05130; A05130.
 DR PIR: S14624; S14624.
 DR PDB: 2CHB; 03-DEC-97.
 DR PDB: 3CHB; 12-AUG-98.
 DR PDB: 1CHP; 08-MAR-96.
 DR PDB: 1CHO; 08-MAR-96.
 DR PDB: 1FGB; 23-DEC-96.
 DR PDB: 1XPB; 01-APR-97.
 DR PDB: 1XTC; 01-AUG-96.

DR PDB: 1CT1; 15-OCT-97.
 DR TIGR: VC1456; -
 DR InterPro: IPR001835; Enterotoxin_B.
 DR Pfam: PF01376; Enterotoxin_B; 1.
 DR PRINTS: PR00772; ENTEROTOXINB.
 DR ProDom: PD012805; Enterotoxin_B; 1.
 KW Membrane; Enterotoxin; Signal; 3D-structure; Complete proteome.
 FT SIGNAL 1 21
 FT CHAIN 22 124
 FT DISULFID 30 107
 FT CONFLICT 33 33
 FT CONFLICT 39 39
 FT CONFLICT 43 43
 FT CONFLICT 68 68
 FT CONFLICT 70 70
 FT CONFLICT 75 75
 FT CONFLICT 91 91
 FT HELIX 26 30
 FT TURN 31 32
 FT TURN 34 35
 FT STRAND 36 44
 FT STRAND 47 51
 FT TURN 54 55
 FT STRAND 58 62
 FT TURN 64 65
 FT STRAND 68 71
 FT TURN 76 77
 FT HELIX 80 99
 FT TURN 100 100
 FT STRAND 102 109
 FT STRAND 115 123
 SQ SEQUENCE 124 AA; 13957 MW; 9AA393E3EA8E3EBF CRC64;
 Query Match 100.0%; Score 39; DB 1; Length 124;
 Best Local Similarity 100.0%; Pred No. 0.23;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EYVPSQH 7
 Db 72 EYVPSQH 78

RESULT 2
 ELBH.ECOLI
 ID ELBH.ECOLI STANDARD; PRT: 124 AA.
 AC P13811;
 DT 01-JAN-1990 (rel. 13, Created)
 DT 01-JAN-1990 (rel. 13, Last sequence update)
 DT 15-JUL-1999 (rel. 38, Last annotation update)
 DE Heat-labile enterotoxin B chain precursor (LT-B, human) (LTH-B).
 GN ELTB OR LTBP.
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Escherichia.
 OX NCBI_TaxID=562;
 OX [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ISOLATE H10407;
 RX MEDLINE=83114628; PubMed=6759877;
 RA Yamamoto T., Tamura T.A., Yokota T., Takano T.;
 RT "Nucleotide sequence comparison between heat-labile toxin B-subunit
 RT cistrons from Escherichia coli of human and porcine origin.";
 RL Infect. Immun. 48:73-77(1985).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ISOLATE H10407;
 RX MEDLINE=83114628; PubMed=6759877;
 RA Yamamoto T., Tamura T.A., Yokota T., Takano T.;
 RT "Overlapping genes in the heat-labile enterotoxin operon originating
 RT from Escherichia coli human strain.";
 RL Mol. Gen. Genet. 188:356-359(1982).
 RN [3]

RP SEQUENCE FROM N.A.
 RC STRAIN-ISOLATE H10407;
 RA MEDLINE=93252225; PubMed=8486242;
 RT Inoue T., Tsuji T., Koto M., Imamura S., Miyama A.;
 RT "Amino acid sequence of heat-labile enterotoxin from chicken
 RT enterotoxigenic Escherichia coli is identical to that of human strain
 RT H 10407.";
 RL FEMS Microbiol. Lett. 108:157-161(1993).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ETEC LT 87;
 RA Germani Y., Desperlier J.M.;
 RL Submitted (OCT-1996) to the EMBL/genbank/DBJ databases.
 RN [5]
 RP DISCUSSION OF SEQUENCE.
 RX MEDLINE=95349400; PubMed=7623669;
 RA Domenighini M., Pizsa M., Jobling M.G., Holmes R.K., Rappuoli R.;
 RT "Identification of errors among database sequence entries and
 RT comparison of correct amino acid sequences for the heat-labile
 RT enterotoxins of Escherichia coli and Vibrio cholerae.";
 RL Mol. Microbiol. 15:1165-1167(1995).
 RN [6]
 RP X-RAY CRYSTALLOGRAPHY (3.04 ANGSTROMS).
 RX MEDLINE=99185101; PubMed=10085117;
 RA Matkovic-Calogovic D., Loregian A., D'Acunto M.R., Battistutta R.,
 RA Tossi A., Palu G., Zanotti G.;
 RT "Crystal structure of the B subunit of Escherichia coli heat-labile
 RT enterotoxin carrying peptides with anti-herpes simplex virus type 1
 RT activity.";
 RL J. Biol. Chem. 274:8764-8769(1999).
 CC -1- FUNCTION: THE BIOLOGICAL ACTIVITY OF THE TOXIN IS PRODUCED BY
 CC THE A CHAIN, WHICH ACTIVATES INTRACELLULAR ADENYL CYCLASE.
 CC -1- SUBUNIT: HETEROHXAMER OF ONE A CHAIN AND OF FIVE B CHAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M17874; AAA98064.1; -
 DR EMBL: J01646; AAB02982.1; -
 DR EMBL: S60731; AAC60441.1; -
 DR EMBL: X83966; CAA58800.1; -
 DR PDB: 1LTR; 23-MAR-99.
 DR InterPro: IPR001835; Enterotoxin_B.
 DR Pfam: PF01376; Enterotoxin_B; 1.
 DR PRINTS: PR00772; ENTEROTOXINB.
 DR ProDom: PD012805; Enterotoxin_B; 1.
 DR Enterotoxin; Signal; 3D-structure.
 FT SIGNAL 1 21
 FT CHAIN 22 124 HEAT-LABILE ENTEROTOXIN B CHAIN.
 FT DISULFID 30 107
 FT VARIANT 6 6 F -> C (IN ISOLATE H10407).
 FT VARIANT 20 20 H -> Y (IN ISOLATE H10407).
 FT VARIANT 34 34 H -> R (IN ISOLATE H10407).
 SO SEQUENCE 124 AA; 14027 MW; E9F7FC7B9D3BC47 CRC64;

Query Match 100.0%; Score 39; DB 1; Length 124;
 Best Local Similarity 100.0%; Pred. No. 0.23;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EVPGSOH 7
 DB 72 EVPGSOH 78

RESULT 3
 EMBL_ECOLI STANDARD; PRT; 124 AA.

AC P32890; P13768; P01557;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)
 GN Heat-labile enterotoxin B chain precursor (LT-B, porcine) (LTP-B).
 GN ELPB OR LTPB.
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Escherichia.
 OX NCBI_TaxId=562;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ISOLATE P307;
 RX MEDLINE=81074965; PubMed=7003397;
 RA Dallas W.S., Falkow S.;
 RT "Amino acid sequence homology between cholera toxin and Escherichia
 RT coli heat-labile toxin.";
 RL Nature 288:499-501(1980).
 RN [2]
 RP REVISIONS TO 28 AND 64.
 RC STRAIN-ISOLATE P307;
 RX MEDLINE=85156481; PubMed=3884513;
 RA Leong J., Vinal A.C., Dallas W.S.;
 RT "Nucleotide sequence comparison between heat-labile toxin B-subunit
 RT cistrons from Escherichia coli of human and porcine origin.";
 RL Infect. Immun. 48:73-77(1985).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ISOLATE PCG86;
 RX MEDLINE=87137303; PubMed=3546273;
 RA Yamamoto T., Gotohori T., Yokota T.;
 RT "Evolutionary origin of pathogenic determinants in enterotoxigenic
 RT Escherichia coli and Vibrio cholerae O1.";
 RL J. Bacteriol. 169:1352-1357(1987).
 RN [4]
 RP SEQUENCE OF 1-22 FROM N.A.
 RX MEDLINE=87280041; PubMed=3301830;
 RA Ibrahim T., Gentz R.;
 RT "A functional interaction between the signal peptide and the
 RT translation apparatus is detected by the use of a single point
 RT mutation which blocks translocation across mammalian endoplasmic
 RT reticulum.";
 RL J. Biol. Chem. 262:10189-10194(1987).
 RN [5]
 RP X-RAY CRYSTALLOGRAPHY (1.95 ANGSTROMS).
 RX MEDLINE=93240541; PubMed=8478941;
 RA Sijma T.K., van Zanten B.A.M., Dauter Z., Hol W.G.J.;
 RT "Refined structure of Escherichia coli heat-labile enterotoxin, a
 RT close relative of cholera toxin.";
 RL J. Mol. Biol. 230:890-918(1993).
 RN [6]
 RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS).
 RX MEDLINE=91238966; PubMed=2034287;
 RA Sijma T.K., Pronk S.E., Kalk K.H., Martna E.S., van Zanten B.A.M.,
 RA Witolt B., Hol W.G.J.;
 RT "Crystal structure of a cholera toxin-related heat-labile enterotoxin
 RT from E. coli.";
 RL Nature 351:371-377(1991).
 RN [7]
 RP DISCUSSION OF SEQUENCE.
 RX MEDLINE=95349400; PubMed=7623669;
 RA Domenighini M., Pizsa M., Jobling M.G., Holmes R.K., Rappuoli R.;
 RT "Identification of errors among database sequence entries and
 RT comparison of correct amino acid sequences for the heat-labile
 RT enterotoxins of Escherichia coli and Vibrio cholerae.";
 RL Mol. Microbiol. 15:1165-1167(1995).
 CC -1- FUNCTION: THE BIOLOGICAL ACTIVITY OF THE TOXIN IS PRODUCED BY
 CC THE A CHAIN, WHICH ACTIVATES INTRACELLULAR ADENYL CYCLASE.
 CC -1- SUBUNIT: HETEROHXAMER OF ONE A CHAIN AND OF FIVE B CHAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; M17873; AAA98065.1; -;
DR EMBL; M15363; AAA24792.1; -;
DR EMBL; M17101; AAA23973.1; -;
DR PIR; A01820; QLECEB.
DR PIR; B26946; QLECEB.
DR PDB; 1LTA; 31-JAN-94.
DR PDB; 1LTA; 31-JAN-94.
DR PDB; 1LTA; 31-JAN-94.
DR PDB; 1LTA; 15-SEP-95.
DR PDB; 1LTA; 17-AUG-96.
DR PDB; 1LTA; 31-JAN-94.
DR PDB; 1LTA; 31-JAN-94.
DR PDB; 1LTA; 07-JUL-97.
DR PDB; 1LTA; 16-JUN-97.
DR PDB; 1LTA; 03-DEC-97.
DR PDB; 1LTA; 03-DEC-97.
DR PDB; 1LTA; 20-APR-95.
DR InterPro; IPR001835; Enterotoxin_B.
DR Pfam; PF01376; Enterotoxin_B; 1.
DR ProDom; PD012805; Enterotoxin_B; 1.
DR Enterotoxin; Signal; 3D-structure.
KW SIGNAL
FT CHAIN 1 21
FT DISULFID 22 124 HEAT-LABILE ENTEROTOXIN B CHAIN.
FT HELIX 26 30
FT TURN 31 32
FT STRAND 36 43
FT TURN 47 51
FT STRAND 54 55
FT TURN 58 62
FT STRAND 64 65
FT TURN 68 71
FT STRAND 76 77
FT TURN 76 77
FT HELIX 80 98
FT TURN 99 100
FT STRAND 103 109
FT STRAND 115 123
SO SEQUENCE 124 AA; 14133 MW; 6DB7DE5835EA70D CRC64;

Query Match 100.0%; Score 39; DB 1; Length 124;
Best Local Similarity 100.0%; Pred. No. 0.23;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EVPGSQH 7
Db 72 EVPGSQH 78

RESULT 4
GUDH_BACSU STANDARD; PRT; 455 AA.
ID GUDH_BACSU
AC P42238;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Probable glucarate dehydratase (EC 4.2.1.40) (GUDH) (glucD).
GN GUDH.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=95219079; PubMed=7704254;
RA Ogawa K.-I., Akagawa E., Nakamura K., Yamane K.;
RT "Determination of a 21548 bp nucleotide sequence around the 24

RT degrees region of the Bacillus subtilis chromosome.";
RL Microbiology 141:269-275(1995).
CC -1- FUNCTION: CATALYZES THE DEHYDRATION OF GLUCARATE TO 5-KETO-4-
CC DEOXY-D-GLUCARATE (5-KDGLUC) (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: D-glucarate + 5-dehydro-4-deoxy-D-glucarate +
CC H₂O.
CC -1- PATHWAY: FIRST STEP IN GLUCARATE CATABOLISM.
CC -1- SIMILARITY: BELONGS TO THE MANDATELASE / MUONATE
CC LACTONIZING ENZYME FAMILY. GLUCD SUBFAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; D30808; BAA06470.1; -;
DR EMBL; Z99105; CAB12043.1; -;
DR HSSP; P42206; 1BQG.
DR Subtilist; BG11161; gudD.
DR InterPro; IPR001354; MR_MLE.
DR Pfam; PF01188; MR_MLE; 1.
DR Pfam; PF02746; MR_MLE_N; 1.
DR Lyase; Complete proteome.
KW SEQUENCE 455 AA; 50782 MW; 3238486007698C2A CRC64;

Query Match 82.1%; Score 32; DB 1; Length 455;
Best Local Similarity 71.4%; Pred. No. 27;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 EVPGSQH 7
Db 64 EVPGSQH 70

RESULT 5
YOG1_CAEEL STANDARD; PRT; 534 AA.
ID YOG1_CAEEL
AC P34610;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Putative serine protease ZK112.1 precursor (EC 3.4.-.-).
GN ZK112.1.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RX MEDLINE=94150718; PubMed=7906398;
RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,
RA Bonfield J., Burton J., Connell M., Cosey T., Cooper J., Coulson A.,
RA Craxton M., Dear S., Du Z., Durbin R., Favell A., Fraser A.,
RA Johnston L., Gardner A., Green P., Hawkins T., Hillier L., Jier M.,
RA Johnston L., Jones M., Kershaw J., Kirsten J., Laister N.,
RA Latreille P., Lightning J., Lloyd C., Mortimore B., O'Callaghan M.,
RA Parsons J., Percy C., Rifkin L., Roopra A., Saunders D., Showkhen R.,
RA Sims M., Smaildon N., Smith A., Smith M., Sonhammer E., Staden R.,
RA Sulston J., Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K.,
RA Waterston R., Watson A., Weinstock L., Wilkinson-Sproat J.,
RA Wooldman P.;
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT elegans.";
RL Nature 368:32-38(1994).
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S28 (SERINE PROTEASE).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

DR EMBL: L14324; AAA28181.1; -
 DR PIR: S44886; S44886.
 DR WormPep: ZK112.1; CE00372.
 DR InterPro: IPR000073; Abhydrolase.
 DR Pfam: PF00561; abhydrolase_1.
 DR Hypothetical protein: Hydrolase; Serine protease; Signal.
 FT SIGNAL 1 17 POTENTIAL.
 FT CHAIN 18 334 POTATIVE SERINE PROTEASE ZK112.1.
 FT ACT_SITE 177 177 CHARGE RELAY SYSTEM (POTENTIAL).
 FT ACT_SITE 420 420 CHARGE RELAY SYSTEM (POTENTIAL).
 FT ACT_SITE 448 448 CHARGE RELAY SYSTEM (POTENTIAL).
 FT CARBOHYD 69 69 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 107 107 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 126 126 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 240 240 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 244 244 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 257 257 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 271 271 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 288 288 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 316 316 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 534 AA; 59221 MW; C43F0104B42E4630 CRC64;

Query Match 82.1%; Score 32; DB 1; Length 534;
 Best Local Similarity 71.4%; Pred. No. 32;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 EYFGSOH 7
 I:| | | |
 Db 442 EIRGSAH 448

RESULT 6
 PD1A_MOUSE STANDARD; PRT; 666 AA.
 AC 092183;
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DE Protein-arginine deiminase type IV (EC 3.5.3.15) (Peptidylarginine deiminase IV).
 GN PD1A.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Epidermis; PubMed=10092850;
 RX MEDLINE=99192810; PubMed=10092850;
 RA Rusd A.A., Ikejiri Y., Ono H., Yonekawa T., Shiraiwa M., Kawada A., Takahara H.;
 RT "Molecular cloning of cDNAs of mouse peptidylarginine deiminase type I, type III and type IV, and the expression pattern of type I in mouse.";
 RT Eur. J. Biochem. 259:660-669(1999).
 CC -1- FUNCTION: CATALYZES THE DEIMINATION OF ARGININE RESIDUES OF PROTEINS.
 CC -1- CATALYTIC ACTIVITY: Protein L-arginine + H(2)O = protein L-citrulline + NH(3).
 CC -1- COFACTOR: REQUIRES CALCIUM IONS.
 CC -1- SIMILARITY: SOME SIMILARITY WITH ARGINASE AND TO ARGININE DEIMINASE.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation -

the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

DR EMBL: AB013850; BAA34246.1; -
 DR MGD: MGI:1338988; PD1A.
 DR InterPro: IPR004303; PAD.
 DR Pfam: PF03068; PAD; 1.
 DR Hydrolase; Calcium-binding; Multigene family.
 FT CA_BIND 505 516 EF-HAND (POTENTIAL).
 SQ SEQUENCE 666 AA; 74476 MW; 70FAE4E7E232D34A CRC64;

Query Match 82.1%; Score 32; DB 1; Length 666;
 Best Local Similarity 71.4%; Pred. No. 40;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 EYFGSOH 7
 I:| | | |
 Db 239 ELPGQGH 245

RESULT 7
 SIM1_MOUSE STANDARD; PRT; 765 AA.
 AC 061045; P70183;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Single-minded homolog 1 (SIM1).
 GN SIM1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SWISS WEBSTER;
 RX MEDLINE=97020303; PubMed=8812055;
 RA Fan C.-M., Kuwana E., Bulfone A., Fletcher C.F., Copeland N.G., Jenkins N.A., Crews S., Martinez S., Puellies L., Rubenstein J.L., Tessier-Lavigne M.;
 RT "Expression patterns of two murine homologs of Drosophila single-minded suggest possible roles in embryonic patterning and in the pathogenesis of Down syndrome.";
 RT Mol. Cell. Neurosci. 7:1-16(1996).
 RN [2]
 RP ERRATUM.
 RX MEDLINE=97029422; PubMed=8875433;
 RA Fan C.-M., Kuwana E., Bulfone A., Fletcher C.F., Copeland N.G., Jenkins N.A., Crews S., Martinez S., Puellies L., Rubenstein J.L., Tessier-Lavigne M.;
 RT Mol. Cell. Neurosci. 7:519-519(1996).
 RN [3]
 RP REVISIONS TO C-TERMINUS.
 RX MEDLINE=97343329; PubMed=9199934;
 RA Fan C.-M.;
 RT Unpublished results, cited by:
 RT Christ R., Scott H.S., Chen H., Kudoh J., Rossier C., Minoshima S., Wang Y., Shimizu N., Antonarakis S.E.;
 RT Genome Res. 7:615-624(1997).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J;
 RX MEDLINE=96413339; PubMed=8927054;
 RA Ena M., Morita M., Ikawa S., Tanaka M., Matsuda Y., Gotoh O., Saijoh Y., Fujii H., Hamada H., Kikuchi Y., Fujii-Kuriyama Y.;
 RT "Two new members of the murine Sim gene family are transcriptional repressors and show different expression patterns during mouse embryogenesis.";
 RT Mol. Cell. Biol. 16:5865-5875(1996).

[5]
RN SEQUENCE FROM N.A.
RP STRAIN-129/SV.
RA Hosoya T.;
RL Submitted (APR-1998) to the EMBL/Genbank/DBJ databases.
RN [6]
RP SUBUNIT.
RX MEDLINE=97172525; PubMed=9020169;
RA Probst M.R., Fan C.M., Tessier-Lavigne M., Hankinson O.;
RT "Two murine homologs of the Drosophila single-minded protein that
interact with the mouse aryl hydrocarbon receptor nuclear
translocator protein."
RL J. Biol. Chem. 272:4451-4457(1997).
CC -1 FUNCTION: TRANSCRIPTIONAL FACTOR THAT MAY HAVE PLEIOTROPIC EFFECTS
DURING EMBRYOGENESIS AND IN THE ADULT.
CC -1 SUBUNIT: EFFICIENT DNA BINDING REQUIRES DIMERIZATION WITH ANOTHER
BHLH PROTEIN. HETERODIMER OF SIM1 AND ARNT.
CC -1 TISSUE SPECIFICITY: DETECTED IN LUNG, SKELETAL MUSCLE AND KIDNEY.
CC DURING FETAL DEVELOPMENT IT IS FOUND IN THE CNS, DEVELOPING
KIDNEY, MESODERMAL AND ENDODERMAL TISSUES, INCLUDING DEVELOPING
SOMITES, MESONEPHRIC DUCT, AND FOREGUT.
CC -1 SIMILARITY: BELONGS TO THE BASIC HELIX-LOOP-HELIX (BHLH) FAMILY OF
TRANSCRIPTION FACTORS.
CC -1 SIMILARITY: CONTAINS 2 PAS (PER-ARNT-SIM) DIMERIZATION DOMAINS.
CC -1 SIMILARITY: CONTAINS 1 PAS-ASSOCIATED C-TERMINAL (PAC) DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U40575; AAA91201.1; ALT_SEQ.
DR EMBL: D79209; BAA11467.1; -
DR EMBL: AB013491; BAA28270.1; -
DR EMBL: AB013484; BAA28270.1; JOINED.
DR EMBL: AB013485; BAA28270.1; JOINED.
DR EMBL: AB013486; BAA28270.1; JOINED.
DR EMBL: AB013487; BAA28270.1; JOINED.
DR EMBL: AB013488; BAA28270.1; JOINED.
DR EMBL: AB013489; BAA28270.1; JOINED.
DR EMBL: AB013490; BAA28270.1; JOINED.
DR MGD: MGI:98306; Sim1.
DR InterPro: IPR003015; HLH_MyC.
DR InterPro: IPR001092; HLH_dlm.
DR InterPro: IPR001067; Nucleinslocator.
DR InterPro: IPR001610; PAC.
DR InterPro: IPR000014; PAS.
DR Pfam: PF00785; PAC; 1.
DR Pfam: PF00989; PAS; 2.
DR PRINTS: PR00785; NCTRNLOCATR.
DR SMART: SM00353; HLH; 1.
DR SMART: SM00086; PAC; 1.
DR SMART: SM00091; PAS; 2.
DR PROSITE: PS00038; HELIX_LOOP_HELIX; 1.
DR PROSITE: PS50112; PAS; 2.
KW Developmental protein; Neurogenesis; Nuclear protein; Repeat;
KW Transcription regulation; DNA-binding.
FT DNA_BIND 1 13 BASIC DOMAIN.
FT DOMAIN 14 54 HELIX-LOOP-HELIX MOTIF (BY SIMILARITY).
FT DOMAIN 77 147 PAS 1.
FT DOMAIN 218 288 PAS 2.
FT CONFLICT 133 133 H -> L (IN REF. 1).
FT CONFLICT 176 176 MISSING (IN REF. 1).
FT CONFLICT 322 322 P -> R (IN REF. 1).
FT CONFLICT 480 480 A -> P (IN REF. 1).
FT CONFLICT 537 537 D -> S (IN REF. 1).
SQ SEQUENCE 765 AA; 85540 MW; B1A7F7DA8578CD17 CRC64;

Query Match

82.1%; Score 32; DB 1; Length 765;

Best Local Similarity 85.7%; Pred. No. 47;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 EVPGSOH 7
DB 423 EVPGSOH 429
RESULT 8
YH96_MYCLE
ID YH96_MYCLE STANDARD; PRT; 137 AA.
AC P13333;
DT 01-JAN-1990 (Rel. 13, Created)
DR 01-JAN-1990 (Rel. 13, Last sequence update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical 15.2 kDa protein ML1796.
GN ML1796.
OS Mycobacterium leprae.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1769;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=8608878; PubMed=2447183;
RA Booth R.J., Harris D.P., Love J.M., Watson J.D.;
RT "Antigenic proteins of Mycobacterium leprae. Complete sequence of the
gene for the 18-kDa protein."
RL J. Immunol. 140:597-601(1988).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-TN.
RX MEDLINE=21128732; PubMed=11234002;
RA Cole S.T., Elgimeier K., Parkhill J., James K.D., Thomson N.R.,
RA Wheeler P.R., Honore N., Garnier T., Churcher C., Harris D.,
RA Mungall K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R.M., Devlin K., Duthoy S., Feltham T., Fraser A., Hamlin N.,
RA Holtroyd S., Hornsby T., Jagels K., Lacroix C., Maclean J., Moule S.,
RA Murphy L., Oliver K., Quail M.A., Rajandream M.A., Rutherford K.M.,
RA Rutter S., Seeger K., Simon S., Simmonds M., Skelton J., Squares R.,
RA Squares S., Stevens K., Taylor K., Whitehead S., Woodward J.R.,
RA Barrell B.G.;
RT "Massive gene decay in the leprosy bacillus."
RL Nature 409:1007-1011(2001).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M19058; AAA88230.1; -
DR EMBL: AL583923; CAC30749.1; -
DR PIR: B27386; B27586.
DR Leprona; ML1796; -
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 137 AA; 15195 MW; 633E68BF3FDAD0DB CRC64;

Query Match
Best Local Similarity 79.5%; Score 31; DB 1; Length 137;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 VPGSOH 7
DB 127 VPGNOH 132

RESULT 9
CIDA_HUMAN
ID CIDA_HUMAN STANDARD; PRT; 219 AA.
AC 060543;

DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Cell death activator CIDF-A (Cell death-inducing DFFA-like effector
 A).
 GN CIDF-A.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP MEDLINE=98232498; PubMed=9564035;
 RA Itohara N., Koseki T., Chen S., Wu X., Nunez G.;
 RT "CIDF, a novel family of cell death activators with homology to the 45
 kDa subunit of the DNA fragmentation factor.";
 RL EMBO J. 17:2526-2533(1998).
 CC -1- FUNCTION: ACTIVATES APOPTOSIS.
 CC -1- SUBUNIT: INHIBITED BY DFFB.
 CC -1- SIMILARITY: CONTAINS 1 CIDF-N DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AF041378; AAC34987.1; -;
 DR HSSP: O9UHD4; 1D4B.
 DR TIM: 604440; -;
 DR InterPro: IPR003508; CAD.
 DR Pfam: PF02017; CIDF-N; 1.
 DR SMART: SM00266; CAD; 1.
 DR APOPTOSIS.
 KW DOMAIN
 FT 33 110 CIDF-N.
 SQ SEQUENCE 219 AA; 24686 MW; 05F704823CE71C0E CRC64;
 Query Match 79.5%; Score 31; DB 1; Length 219;
 Best Local Similarity 83.3%; Pred. No. 20;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 2 VPGSQH 7
 DB 109 MPGSQH 114
 RESULT 10
 ID YBSA_THEME STANDARD; PRT; 240 AA.
 AC P58009;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical protein TM1158.1.
 GN TM1158.1.
 OS Thermotoga maritima.
 OC Bacteria; Thermotogales; Thermotoga.
 OX NCBI_TaxID=2336;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MSB8 / DSM 3109;
 RX MEDLINE=99287316; PubMed=10360571;
 RA Nelson K.E., Clayton R.A., Gill S.R., Gwin M.L., Dodson R.J.,
 RA Haft D.H., Hickey E.K., Peterson J.D., Nelson W.C., Ketchum K.A.,
 RA McDonald L., Uitterback T.R., Malek J.A., Linher K.D., Garrett M.M.,
 RA Stewart A.M., Cotton M.D., Pratt M.S., Phillips C.A., Richardson D.,
 RA Heidelberg J., Sutton G.G., Fleischmann R.D., Eisen J.A., White O.,
 RA Salzberg S.L., Smith H.O., Venter J.C., Fraser C.M.;
 RT "Evidence for lateral gene transfer between Archaea and Bacteria from
 genome sequence of Thermotoga maritima.";

RL Nature 399:323-329(1999).
 RN [2]
 RP IDENTIFICATION.
 RA Medigue C., Bocs S.;
 RL Unpublished Observations (APR-2001).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AE001773; -; NOT_ANNOTATED.CDS.
 DR TIGR: TM1158.1; -;
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 240 AA; 27773 MW; 1BEF66C1C8BD2700 CRC64;
 Query Match 79.5%; Score 31; DB 1; Length 240;
 Best Local Similarity 83.3%; Pred. No. 22;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 2 VPGSQH 7
 DB 226 VPGSEH 231
 RESULT 11
 ID DLX3_NOTVTI STANDARD; PRT; 273 AA.
 AC P53770;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE Homeobox protein DLX-3 (Box-4) (NVHBOX-4).
 GN DLX3 OR BOX4.
 OS Notoptentulus viridescens (Eastern newt) (Triturus viridescens).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Amphibia; Batrachia; Caudata; Salamandridae; Salamandridae;
 CC Notoptentulus.
 OX NCBI_TaxID=8316;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Skin;
 RX MEDLINE=93050784; PubMed=1358728;
 RA Beauchemin M., Savard P.;
 RT "Two distal-less related homeobox-containing genes expressed in
 regeneration blastemas of the newt.";
 RL Dev. Biol. 154:55-65(1992).
 CC -1- SUBCELLULAR LOCATION: Nuclear (Potential).
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN REGENERATION BLASTEMAS. DETECTED
 CC IN FORELIMBS, HINDLIMBS, THE TAIL, FLANK, AND BRAIN AS WELL AS IN
 CC LIMB AND TAIL BLASTEMAS.
 CC -1- SIMILARITY: BELONGS TO THE DISTAL-LESS FAMILY OF HOMEOBOX
 CC PROTEINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X63531; CAA45094.1; -;
 DR HSSP: P22808; INK3.
 DR InterPro: IPR000047; HTH_repressr.
 DR InterPro: IPR001356; Homeobox.
 DR Pfam: PF00046; homeobox; 1.
 DR PRINTS: PR00024; HOMEOBOX.
 DR PRINTS: PR00031; HTHREPRESSR.

DR SMART: SM00389; HOX; 1.
DR PROSITE: PS00027; HOMEBOX_1; 1.
DR PROSITE: PS50071; HOMEBOX_2; 1.
KW Homeobox; DNA-binding; Developmental protein; Nuclear protein.
FT DNA_BIND 125 184 HOMEBOX.
SQ SEQUENCE 273 AA; 30654 MW; B356D01233061F2F CRC64;

Query Match 79.5%; Score 31; DB 1; Length 273;
Best Local Similarity 71.4%; Pred. No. 26;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 EYVGSQH 7
||||:|
Db 188 EYVGMH 194

RESULT 12
DLX3_PLEWA STANDARD; PRT; 274 AA.

AC Q91284;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Distal-less like protein DLX-3.
GN DLX3.

OS Pleurodeles waltl (Iberian ribbed newt).
OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Caudata; Salamandroidae; Salamandridae;
OC Pleurodeles.

OX NCBI_TaxID=8319;

RN [1]

RP SEQUENCE FROM N.A.
RX MEDLINE-96391186; PubMed-8798159;

RA Nicolas S., Massacrier A., Caubit X., Cau P., le Parco Y.;
RT "A distal-less-like gene is induced in the regenerating central
nervous system of the urodele Pleurodeles waltl.";

RL Mech. Dev. 56:209-220(1996).

CC -1- SUBCELLULAR LOCATION: Nuclear (Potential).
CC -1- DEVELOPMENTAL STAGE: EXPRESSED IN THE REGENERATING SPINAL CORD BUT
NOT IN THE ADULT ONE.

CC -1- SIMILARITY: BELONGS TO THE DISTAL-LESS FAMILY OF HOMEBOX
PROTEINS.

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@sib-sib.ch).

CC EMBL: U49645; AAA98645.1; -
DR HSSP: P22808; INK3.

DR InterPro: IPR000047; HTH_repressr.
DR InterPro: IPR001356; Homeobox.

DR Pfam: PF00046; homeobox.1.

DR PRINTS: PR00024; HOMEBOX.
DR PRINTS: PR00031; HTHREPRESSR.

DR SMART: SM00389; HOX; 1.

DR PROSITE: PS00027; HOMEBOX_1; 1.
DR PROSITE: PS50071; HOMEBOX_2; 1.

DR Homeobox; DNA-binding; Developmental protein; Nuclear protein.
KW DNA_BIND 126 185 HOMEBOX.
SQ SEQUENCE 274 AA; 30607 MW; 94148553BA808C8 CRC64;

Query Match 79.5%; Score 31; DB 1; Length 274;
Best Local Similarity 71.4%; Pred. No. 26;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 EYVGSQH 7
||||:|

Db 189 EYVGMH 195

RESULT 13
DLX3_AMBE STANDARD; PRT; 280 AA.

AC Q90229;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Homeobox protein DLX-3.
GN DLX-3.

OS Ambystoma mexicanum (Axolotl).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Caudata; Salamandroidae; Ambystomatidae;
OC Ambystoma.

OX NCBI_TaxID=8296;

RN [1]

RP SEQUENCE FROM N.A.
RX MEDLINE-97108743; PubMed-8951064;

RA Mullen L.M., Bryant S.V., Torok M.A., Blumberg B., Gardiner D.M.;
RT "Nerve dependency of regeneration: the role of Distal-less and FGF
signaling in amphibian limb regeneration.";

RL Development 122:3487-3497(1996).

CC -1- SUBCELLULAR LOCATION: Nuclear (Potential).
CC -1- SIMILARITY: BELONGS TO THE DISTAL-LESS FAMILY OF HOMEBOX
PROTEINS.

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@sib-sib.ch).

CC EMBL: U59480; AAB49668.1; -

DR HSSP: P22808; INK3.
DR InterPro: IPR000047; HTH_repressr.

DR InterPro: IPR001356; Homeobox.
DR Pfam: PF00046; homeobox.1.

DR PRINTS: PR00024; HOMEBOX.
DR PRINTS: PR00031; HTHREPRESSR.

DR SMART: SM00389; HOX; 1.

DR PROSITE: PS00027; HOMEBOX_1; 1.
DR PROSITE: PS50071; HOMEBOX_2; 1.

DR Homeobox; DNA-binding; Developmental protein; Nuclear protein.
KW DNA_BIND 126 185 HOMEBOX.
SQ SEQUENCE 280 AA; 31200 MW; 0F8183097ABAC791 CRC64;

Query Match 79.5%; Score 31; DB 1; Length 280;
Best Local Similarity 71.4%; Pred. No. 26;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 EYVGSQH 7
||||:|
Db 189 EYVGMH 195

RESULT 14
COT2_BOVIN STANDARD; PRT; 414 AA.

AC Q9TR7;
DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE COUP transcription factor 2 (COUP-TF2) (COUP-TF II).
GN NR2P2 OR TFCOUP2.

OS Bos taurus (Bovine).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;

OC Bovidae; Bovinae; Bos.

OX NCBI_TaxID=9913;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Corpus luteum;
 RA Walther N.;
 RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: REGULATION OF THE APOLIPOPROTEIN AI GENE TRANSCRIPTION.
 CC BINDS TO DNA SITE A (BY SIMILARITY).
 CC -1- SUBUNIT: BINDS DNA AS AN HOMODIMER (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC -1- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTORS FAMILY.
 CC NR2 SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AJ249441; CAB5624.1; -.
 CC HSSP: P19793; 2NLL.
 CC InterPro: IPR000536; Hormone_rec.1lg.
 CC InterPro: IPR001723; Steroidhormone_receptor.
 CC InterPro: IPR001628; zf-C4.
 CC Pfam: PF00104; hormone_rec.1.
 CC Pfam: PF00105; zf-C4; 1.
 CC PRINTS: PR00398; STRODHOMNER.
 CC PRINTS: PR00047; STROIDFINGER.
 CC SMART: SM00430; HOL1.1.
 CC SMART: SM00399; ZNF_C4; 1.
 CC PROSITE: PS00031; NUCLEAR_RECEPTOR; 1.
 CC Transcription regulation; DNA-binding; Nuclear protein;
 CC Zinc-finger.
 CC KW DOMAIN 71 75 POLY-GLN.
 CC FT DNA_BIND 79 144 NUCLEAR RECEPTOR-TYPE.
 CC FT ZN_FING 79 99 C4-TYPE.
 CC FT ZN_FING 115 139 C4-TYPE.
 CC FT DOMAIN 337 414 IMPORTANT FOR DIMERIZATION (BY
 CC SIMILARITY).
 CC SQ SEQUENCE 414 AA; 4557 MW; C24CB023C8A27F57 CRC64;
 Query Match 79.5%; Score 31; DB 1; Length 414;
 Best Local Similarity 100.0%; Pred. No. 40;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 EVPGSQ 6
 Db 14 EVPGSQ 19
 RESULT 15
 COT2_HUMAN
 ID COT2_HUMAN STANDARD; PRT; 414 AA.
 AC P24468; Q03754;
 DT 01-MAR-1992 (Rel. 21, Created)
 DT 01-MAR-1992 (Rel. 21, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE COUP transcription factor 2 (COUP-TF2) (COUP-TF II) (Apolipoprotein AI
 DE regulatory protein-1) (ARP-1).
 GN NR2F2 OR TRFCOUP2 OR ARP1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91118002; PubMed=1899293;
 RA Ladias J.A.A., Karathanasis S.K.;
 RT "Regulation of the apolipoprotein AI gene by ARP-1, a novel member of
 the steroid receptor superfamily."

RL Science 251:561-565(1991).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Speckmayer R.W.M., Paulweber B., Sandhofer F.;
 RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE OF 1-351 FROM N.A.
 RX MEDLINE=92314709; PubMed=1820218;
 RA Wang L.H., Ing N.H., Tsai S.Y., O'Malley B.W., Tsai M.J.;
 RT "The COUP-TFs compose a family of functionally related transcription
 RT factors."
 RL Gene Expr. 1:207-216(1991).
 CC -1- FUNCTION: REGULATION OF THE APOLIPOPROTEIN AI GENE TRANSCRIPTION.
 CC BINDS TO DNA SITE A.
 CC -1- SUBUNIT: BINDS DNA AS AN HOMODIMER.
 CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC -1- TISSUE SPECIFICITY: FOUND IN ALL TISSUES.
 CC -1- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTORS FAMILY.
 CC NR2 SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: M64497; AAA86429.1; -.
 CC EMBL: U60477; AAB09475.1; -.
 CC EMBL: M62760; AAA21479.1; -.
 CC PIR: A37133; A37133.
 CC HSSP: P19793; 2NLL.
 CC TRANSFAC: T00045; -.
 CC MIM: 107773; -.
 CC InterPro: IPR000536; Hormone_rec.1lg.
 CC InterPro: IPR001723; Steroidhormone_receptor.
 CC InterPro: IPR001628; zf-C4.
 CC Pfam: PF00104; hormone_rec.1.
 CC Pfam: PF00105; zf-C4; 1.
 CC PRINTS: PR00398; STRODHOMNER.
 CC PRINTS: PR00047; STROIDFINGER.
 CC SMART: SM00430; HOL1.1.
 CC SMART: SM00399; ZNF_C4; 1.
 CC PROSITE: PS00031; NUCLEAR_RECEPTOR; 1.
 CC Transcription regulation; DNA-binding; Nuclear protein;
 CC Zinc-finger.
 CC KW DOMAIN 71 75 POLY-GLN.
 CC FT DNA_BIND 79 144 NUCLEAR RECEPTOR-TYPE.
 CC FT ZN_FING 79 99 C4-TYPE.
 CC FT ZN_FING 115 139 C4-TYPE.
 CC FT DOMAIN 337 414 IMPORTANT FOR DIMERIZATION.
 CC SQ SEQUENCE 414 AA; 4557 MW; C24CB2E8C8A27E8C CRC64;
 Query Match 79.5%; Score 31; DB 1; Length 414;
 Best Local Similarity 100.0%; Pred. No. 40;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 EVPGSQ 6
 Db 14 EVPGSQ 19

Search completed: July 3, 2002, 09:04:41
 Job time: 581 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 3, 2002, 08:55:46 ; Search time 74.17 Seconds
(without alignments)
16.327 Million cell updates/sec

Title: US-09-786-648-2
Perfect score: 39
Sequence: 1 EYPSQH 7

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL.19:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phase:*
10: sp_plant:*
11: sp_prodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	39	100.0	103	2	Q9R646
2	39	100.0	124	2	Q57193
3	39	100.0	124	2	Q9R615
4	39	100.0	124	2	Q56635
5	39	100.0	124	2	Q93V32
6	39	100.0	124	9	Q94M01
7	35	89.7	100	10	Q22511
8	34	87.2	195	13	Q9W7D3
9	34	87.2	199	13	Q9W7D4
10	34	87.2	201	2	Q9R3Z8
11	33	84.6	192	4	Q9NP06
12	33	84.6	272	4	Q9UG04
13	33	84.6	367	10	Q80418
14	33	84.6	412	4	Q9Y2B3
15	33	84.6	868	5	Q9VFE2
16	32	82.1	95	10	Q9F018

17	32	82.1	137	16	Q92T78	Q92T78 rhizobium m
18	32	82.1	765	11	Q70284	Q70284 mus musculu
19	31	79.5	95	10	Q9F0J8	Q9F0J8 amarantus
20	31	79.5	122	5	Q9W2V0	Q9W2V0 drosophila
21	31	79.5	134	16	Q92N40	Q92N40 rhizobium m
22	31	79.5	222	11	Q9JL95	Q9JL95 mus musculu
23	31	79.5	233	10	Q94KK7	Q94KK7 arabidopsis
24	31	79.5	242	2	Q86582	Q86582 streptomyce
25	31	79.5	243	11	Q9QY63	Q9QY63 mus musculu
26	31	79.5	260	10	Q9MA16	Q9MA16 arabidopsis
27	31	79.5	274	17	Q9HNM5	Q9HNM5 halobacteri
28	31	79.5	331	5	Q18391	Q18391 drosophila
29	31	79.5	353	16	Q9RVJ7	Q9RVJ7 delnoccocus
30	31	79.5	360	11	Q9B6W8	Q9B6W8 mus musculu
31	31	79.5	371	2	Q9RH15	Q9RH15 zymomonas m
32	31	79.5	374	16	Q9PKY2	Q9PKY2 chlamydia m
33	31	79.5	382	5	Q9V7M7	Q9V7M7 drosophila
34	31	79.5	406	3	Q96V07	Q96V07 cryphonectr
35	31	79.5	432	3	Q9P3J3	Q9P3J3 neurospora
36	31	79.5	439	11	Q9R0C1	Q9R0C1 mus musculu
37	31	79.5	460	10	Q9LIF8	Q9LIF8 arabidopsis
38	31	79.5	472	2	Q939I3	Q939I3 klebsiella
39	31	79.5	489	5	Q9XUM8	Q9XUM8 caenorhabdi
40	31	79.5	496	11	Q9D0H2	Q9D0H2 mus musculu
41	31	79.5	497	4	Q9BRR6	Q9BRR6 homo sapien
42	31	79.5	510	4	Q9H846	Q9H846 homo sapien
43	31	79.5	510	4	Q96GM8	Q96GM8 homo sapien
44	31	79.5	559	6	Q95JP7	Q95JP7 macaca fasc
45	31	79.5	660	10	Q9AY34	Q9AY34 oryza sativ

ALIGNMENTS

RESULT 1
ID Q9R646 PRELIMINARY; PRT; 103 AA.
AC Q9R646;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE CHOLERA-LIKE ENTEROTOXIN B SUBUNIT.
OS Vibrio cholerae.
OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
OX NCBI_TaxID=666;
RN [1]
RP SEQUENCE.
RX MEDLINE=95303036; PubMed=7783690;
RA Nakashima K., Eguichi Y., Nakasone N.;
RT "Characterization of an enterotoxin produced by Vibrio cholerae
RT O139.";
RL Microbiol. Immunol. 39:87-94(1995).
DR HSSP: P01556; IYTC.
DR InterPro: IPR001835; Enterotoxin_B.
DR Pfam: PF01376; Enterotoxin_B.1.
DR PRINTS: PR00772; ENTEROTOXINB.
DR PRODOM: PD012805; Enterotoxin_B.1.
SQ SEQUENCE 103 AA; 11645 MW; 992A05C07F2F70E CRC64;

Query Match 100.0%; Score 39; DB 2; Length 103;
Best Local Similarity 100.0%; Pred. No. 0.53;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EYPSQH 7
Db 51 EYPSQH 57
RESULT 2
ID Q57193 PRELIMINARY; PRT; 124 AA.
AC Q57193;

DT 01-NOV-1996 (TREMblrel. 01, Created)
DT 01-NOV-1996 (TREMblrel. 01, last sequence update)
DT 01-JUN-2001 (TREMblrel. 17, last annotation update)
DE CHOLERA TOXIN B PROTEIN (CTB) PRECURSOR (CTB).
GN CTXB.
OS *Vibrio cholerae*.
OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; *Vibrio*.
OX NCBI_TaxID=666;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-CLASSICAL STRAIN 569B;
RX MEDLINE=91355224; PubMed=1683840;
RA Dams E., De Wolf M., Dietrick W.;
RT "Nucleotide sequence analysis of the CT operon of the *Vibrio cholerae*
RL classical strain 569B.";
RL Biochim. Biophys. Acta 1090:139-141(1991).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-CLASSICAL BIOTYPE 569B;
RA Shi C., Cao C., Zhang J., Ma Q.;
RL Chin. Biochem. J. 9:395-399(1993).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-CLASSICAL BIOTYPE 569B;
RA Xu L.;
RT Submitted (APR-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL: X58785; CA41591.1; -;
DR EMBL: U25679; AAC34728.1; -;
DR EMBL: A00931; CA000098.1; -;
DR HSSP: P01556; 2CHB.
DR InterPro: IPR001835; Enterotoxin_B.
DR Pfam: PF01376; Enterotoxin_B; 1.
DR PRINTS: PR00772; ENTEROTOXINB.
DR ProDom: PD012805; Enterotoxin_B; 1.
KW signal.
FT SIGNAL.
FT CHAIN 1 21 POTENTIAL.
FT SEQUENCE 124 AA; 13919 MW; D6BF83FF7924EA3 CRC64;
SQ

Query Match 100.0%; Score 39; DB 2; Length 124;
Best Local Similarity 100.0%; Pred. No. 0.65;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EVPGSOH 7
DB 72 EVPGSOH 78

RESULT 3
Q9RP15 PRELIMINARY; PRT; 124 AA.
AC Q9RP15;
DT 01-MAY-2000 (TREMblrel. 13, Created)
DT 01-MAY-2000 (TREMblrel. 13, last sequence update)
DT 01-JUN-2001 (TREMblrel. 17, last annotation update)
DE CHOLERA ENTEROTOXIN B-SUBUNIT.
GN CTXB.
OS *Vibrio cholerae*.
OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; *Vibrio*.
OX NCBI_TaxID=666;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-KNIH002;
RA Shin H.J., Park Y.C., Kim Y.C.;
RT "Cloning and nucleotide sequence analysis of the virulence gene
RT cassette from *Vibrio cholerae* KNIH002 isolated in Korea.";
RL Misalimurtag Holji 35:205-210(1999).
DR EMBL: AF175708; AAD51360.1; -;
DR HSSP: P01556; 2CHB.
DR InterPro: IPR001835; Enterotoxin_B.
DR Pfam: PF01376; Enterotoxin_B; 1.
DR PRINTS: PR00772; ENTEROTOXINB.
RN

DR ProDom: PD012805; Enterotoxin_B; 1.
SQ SEQUENCE 124 AA; 13905 MW; 23BF83FF793E5B9 CRC64;

Query Match 100.0%; Score 39; DB 2; Length 124;
Best Local Similarity 100.0%; Pred. No. 0.65;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EVPGSOH 7
DB 72 EVPGSOH 78

RESULT 4
Q56635 PRELIMINARY; PRT; 124 AA.
AC Q56635;
DT 01-NOV-1996 (TREMblrel. 01, Created)
DT 01-NOV-1996 (TREMblrel. 01, last sequence update)
DT 01-JUN-2001 (TREMblrel. 17, last annotation update)
DE CHOLERA TOXIN PRECURSOR.
GN CTXB.
OS *Vibrio cholerae*.
OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; *Vibrio*.
OX NCBI_TaxID=666;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-S7;
RA Yamamoto K., Do V.G., Xu M., Iida T., Miwatani T., Albert M.J.,
RA Honda T.;
RT "Comparison of cholera toxin genes (ctxAB) of non-O1 *Vibrio cholerae*
RT strains 854 (O139-bengal) and S7 (O37) from two outbreaks.";
RL Submitted (APR-1994) to the EMBL/GenBank/DBJ databases.
DR EMBL: D30052; BAA06289.1; -;
DR HSSP: P01556; 2CHB.
DR InterPro: IPR001835; Enterotoxin_B.
DR Pfam: PF01376; Enterotoxin_B; 1.
DR PRINTS: PR00772; ENTEROTOXINB.
DR ProDom: PD012805; Enterotoxin_B; 1.
KW signal.
FT SIGNAL.
FT SEQUENCE 124 AA; 13871 MW; 3F87B2F297953179 CRC64;
SQ

Query Match 100.0%; Score 39; DB 2; Length 124;
Best Local Similarity 100.0%; Pred. No. 0.65;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EVPGSOH 7
DB 72 EVPGSOH 78

RESULT 5
Q93V32 PRELIMINARY; PRT; 124 AA.
AC Q93V32;
DT 01-DEC-2001 (TREMblrel. 19, Created)
DT 01-DEC-2001 (TREMblrel. 19, last sequence update)
DT 01-DEC-2001 (TREMblrel. 19, last annotation update)
DE HEAT-LABILE ENTEROTOXIN B SUBUNIT.
GN LTH B SUBUNIT.
OS *Escherichia coli*.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-1032 (ENTEROTOXIGENIC);
RA Komase K.;
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.

RC STRAIN-1032 (ENTEROTOXIGENIC);
 RX MEDLINE-95091056; PubMed-7998417;
 RA Tamura S., Asanuma H., Tomita T., Komase K., Kawahara K., Danbara H.,
 RA Hattori N., Watanabe K., Suzuki Y., Nagamine T., Aizawa C., Oya A.,
 RA Kurita T.;
 RT "Escherichia coli heat-labile enterotoxin B subunits supplemented with
 RT a trace amount of the holotoxin as an adjuvant for nasal influenza
 RT vaccine.";
 RL Vaccine 12;1083-1089(1994).
 DR EMBL: AB011677; BAA25726.1; -
 SQ SEQUENCE 124 AA; 14028 MW; 5346BD38B32354C2 CRC64;

Query Match 100.0%; Score 39; DB 2; Length 124;
 Best Local Similarity 100.0%; Pred. No. 0.65;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EYVGSQH 7
 DB 72 EYVGSQH 78

RESULT 6
 O94M01 PRELIMINARY; PRT; 124 AA.
 AC O94M01;
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE CTXB.
 GN CTXB.
 OS Vibrio phage CTX.
 OC Viruses; ssDNA viruses; Inoviridae; Inovirus.
 OX NCBI_TaxID=141904;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Bhattacharyya T., Nandy R.K., Nair G.B.;
 RT "The entire core region of the ctx-phi (ctx-phage) in VCE 232, an
 RT environmental strain of V. cholerae.";
 RL Submitted (Aug-2001) to the EMBL/Genbank/DBJ databases.
 DR EMBL: AF414369; AAL09682.1; -
 SQ SEQUENCE 124 AA; 13884 MW; 8D4C251591B16891 CRC64;

Query Match 100.0%; Score 39; DB 9; Length 124;
 Best Local Similarity 100.0%; Pred. No. 0.65;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EYVGSQH 7
 DB 72 EYVGSQH 78

RESULT 7
 O22511 PRELIMINARY; PRT; 565 AA.
 AC O22511;
 DT 01-JAN-1998 (TREMBLrel. 05, Created)
 DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE GLUTATHIONE REDUCTASE (NADPH) (EC 1.6.4.2) (FRAGMENT).
 GN GOR.
 OS Vitis vinifera (Grape).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Vitaceae;
 OC Vitis.
 OX NCBI_TaxID=29760;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA STRAIN-CV. THOMPSON SEEDLESS (CLONE 2A); TISSUE=FRUIT;
 RA Casool T., Adams D.O.;
 RT "Cloning of a grape glutathione reductase cDNA and analysis of its
 RT expression during berry development.";

RL Submitted (Aug-1997) to the EMBL/Genbank/DBJ databases.
 CC -I- COFACTOR: FAD (BY SIMILARITY).
 DR EMBL: AF019907; AAB70837.1; -
 DR HSSP; P06715; 1GER.
 DR InterPro: IPR001327; FAD_pyr_redox.
 DR InterPro: IPR001033; Pyridine_redox_2.
 DR InterPro: IPR001100; pyr_redox.
 DR InterPro: IPR004099; pyr_redox_dim.
 DR Pfam: PF00070; pyr_redox_1.
 DR Pfam: PF02852; pyr_redox_dim; 1.
 DR PRINTS: PR00368; FADPNR.
 DR PRINTS: PR00411; PNDRFASFI.
 DR PRINTS: PR00469; PNDRFASFI.
 DR PROSITE: PS00076; PYRIDINE_REDOX_1; 1.
 KW FAD; Flavoprotein; Oxidoreductase.
 FT NON_TER
 SQ SEQUENCE 565 AA; 60695 MW; B26113AE09A121DE CRC64;

Query Match 89.7%; Score 35; DB 10; Length 565;
 Best Local Similarity 71.4%; Pred. No. 25;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 EYVGSQH 7
 DB 239 EYVGSQH 245

RESULT 8
 O9W7D3 PRELIMINARY; PRT; 195 AA.
 AC O9W7D3;
 DT 01-NOV-1999 (TREMBLrel. 12, Created)
 DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE HYPOTHETICAL 21.4 KDA PROTEIN.
 OS Oryzias latipes (Medaka fish).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 OC Acanthopterygii; Acanthopterygii; Percomorpha; Atherinomorpha;
 OC Belontiiformes; Adrianichthyidae; Oryziinae; Oryzias.
 OX NCBI_TaxID=8090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ORANGE-RED;
 RX MEDLINE-20070874; PubMed-10602271;
 RA Kanamori A.;
 RT "Systematic identification of genes expressed during early oogenesis
 RT in medaka.";
 RL Mol. Reprod. Dev. 55;31-36(2000).
 DR EMBL: AF12818; AAD38915.1; -
 KW Hypothetical protein.
 SQ SEQUENCE 195 AA; 21446 MW; B90682D695729E88 CRC64;

Query Match 87.2%; Score 34; DB 13; Length 195;
 Best Local Similarity 100.0%; Pred. No. 13;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 VPSQH 7
 DB 111 VPSQH 116

RESULT 9
 O9W7D4 PRELIMINARY; PRT; 199 AA.
 AC O9W7D4;
 DT 01-NOV-1999 (TREMBLrel. 12, Created)
 DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE HYPOTHETICAL 21.7 KDA PROTEIN.
 OS Oryzias latipes (Medaka fish).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorphi; Acanthopterygii; Percomorpha; Atherinomorphi;
OC Belontiiformes; Adrianchthyidae; Oryziatidae; Oryziatidae;
OX NCBI_TaxID=8090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ORANGE-RED;
RA MEDLINE=20070874; PubMed=10602271;
RX Kanamori A.;
RT "Systematic identification of genes expressed during early oogenesis
in medaka.";
RL Mol. Reprod. Dev. 55:31-36(2000).
DR EMBL; AF128817; AAD38914.1; -.
KW Hypothetical protein.
SQ SEQUENCE 199 AA; 21726 MW; E9827C05451B15CD CRC64;

Query Match 87.2%; Score 34; DB 13; Length 199;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 EVPGSOH 7
Db 115 EVPGSOH 120

RESULT 10

O9RJZ8 PRELIMINARY; PRT; 201 AA.

AC O9RJZ8;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE PUTATIVE AMIDASE.
GN SCG11A.03.
OS Streptomyces coelicolor.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2).
RA Oliver K., Harris D.;
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2).
RA Cerdano A.M., Parkhill J., Barrell B.G., Rajandream M.A.;
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2).
RX MEDLINE=97000351; PubMed=8843436;
RA Redenbach M., Krieser H.M., Denapite D., Eichner A., Cullum J.,
RT "A set of ordered cosmid and a detailed genetic and physical map for
the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
RL Mol. Microbiol. 21:77-96(1996).
DR EMBL; AL133210; CAB61584.1; -.
DR InterPro; IPR002502; Amidase_2.
DR Pfam; PF01510; Amidase_2; 1.
SQ SEQUENCE 201 AA; 22749 MW; B8E477E06A20468 CRC64;

Query Match 87.2%; Score 34; DB 2; Length 201;
Best Local Similarity 85.7%; Pred. No. 14;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 EVPGSOH 7
Db 176 EVPGSOH 182

RESULT 11

O9NPQ6 PRELIMINARY; PRT; 192 AA.

ID O9NPQ6;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
DE LPL, LCAT-LIKE LYOPHOSPHOLIPASE (FRAGMENT).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Pluvinet R., Estivill X., Escarceller M., Sumoy L.;
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Aulicay C., Ansoorge W., Ballabio A., Estivill X., Gibson K.,
RA Lehrach H., Poustka A., Lundeberg J.;
RT "The European IMAGE consortium for integrated molecular analysis of
human gene transcripts.";
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL389957; CAB97531.1; -.
FT NON_TER
SQ SEQUENCE 192 AA; 21609 MW; 04A7AE8CB344F213 CRC64;

Query Match 84.6%; Score 33; DB 4; Length 192;
Best Local Similarity 71.4%; Pred. No. 21;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 EVPGSOH 7
Db 166 EVPGSOH 172

RESULT 12

O9UGO4 PRELIMINARY; PRT; 272 AA.

ID O9UGO4;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE HYPOTHETICAL 31.0 KDA PROTEIN (UNKNOWN) (PROTEIN FOR MGC:12657).
GN DKF2P564A0122.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BRIN;
RA Wanduit R., Heubner D., Mewes H.W., Gassenhuber J., Wiemann S.;
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=BRIN, NEUROBLASTOMA;
RA Strausberg R.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL110209; CAB53675.1; -.
DR EMBL; BC011640; AAH1640.1; -.
DR InterPro; IPR003386; LACT.
DR Pfam; PF02450; LACT; 1.
KW Hypothetical protein.
SQ SEQUENCE 272 AA; 31016 MW; ACCCE1680D7A720 CRC64;

Query Match 84.6%; Score 33; DB 4; Length 272;
Best Local Similarity 71.4%; Pred. No. 31;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 EVPGSOH 7
|:|:|:|
Db 246 ELPGSEH 252

RESULT 13
ID 080418 PRELIMINARY; PRT; 367 AA.
AC 080418;
DT 01-NOV-1998 (TREMBLrel. 08, Created)
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DE 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
GN NTC16 PROTEIN.
OS Nicotiana tabacum (Common tobacco).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Solanales; Solanaceae; Nicotiana.
OX NCBI_TaxID=4097;
RN [1]
RP SEQUENCE FROM N.A.
RA Tanaka-Ideguchi M., Itoh H., Oyama N., Koshioka M., Matsuo K.;
RT "Over-expression of a tobacco homeobox gene, NTH15, decreases the
RT expression of a gibberellin biosynthetic gene encoding GA 20-
RT oxidase.";
RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB016084; BAA31690.1; -
DR InterPro: IPR002419; Fe_Asc_oxidored. 1.
DR Pfam: PF00671; Fe_Asc_oxidored; 1.
SQ SEQUENCE 367 AA; 42170 MW; 923BC90B3BBAC05 CRC64;

Query Match 84.6%; Score 33; DB 10; Length 367;
Best Local Similarity 85.7%; Pred. No. 43;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 1 EVPGSOH 7
|:|:|:|
Db 149 EVPGSOH 155

RESULT 14
ID 09Y2B3 PRELIMINARY; PRT; 412 AA.
AC 09Y2B3;
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DE 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE ICAT-LIKE LYSOPHOSPHOLIPASE (LPL).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE-99194552; PubMed-10092508;
RA Tanigawa Y., Shibata S., Kita S., Horikoshi K., Shirafuji H.,
RA Sumino Y., Fujino M.;
RT "Cloning and expression of a novel lysophospholipase which
RT structurally resembles lecithin cholesterol acyltransferase.";
RL Biochem. Biophys. Res. Commun. 257:50-56(1999).
DR EMBL: AB017494; BAA76877.1; -
DR InterPro: IPR003386; LACT.
DR Pfam: PF02450; LACT; 1.
SQ SEQUENCE 412 AA; 46657 MW; 1FEA8A5783AF050A CRC64;

Query Match 84.6%; Score 33; DB 4; Length 412;
Best Local Similarity 71.4%; Pred. No. 49;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
OY 1 EVPGSOH 7
|:|:|:|

Db 386 ELPGSEH 392

RESULT 15
ID 09VEE2 PRELIMINARY; PRT; 868 AA.
AC 09VEE2;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DE 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
GN CG3837 PROTEIN.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BERKELEY;
RX MEDLINE-20196006; PubMed-10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazer V.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA April J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkov D., Botchan M.R., Bouck J., Brokstein P., Brotler P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Dou P.L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Fodor C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Gloder A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Idegawa C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang X., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Mishina N.V., Moberg C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spler E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Sylkes R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
DR EMBL: AE003706; AAF55118.1; -
DR FlyBase: FBgn0038279; CG3837.
DR InterPro: IPR000494; EGFR-L.
DR Pfam: PF01030; Recep_L_domain; 2.
SQ SEQUENCE 868 AA; 98349 MW; F6562A64E72E7B21 CRC64;

Query Match 84.6%; Score 33; DB 5; Length 868;
Best Local Similarity 71.4%; Pred. No. 11e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
OY 1 EVPGSOH 7
|:|:|:|
Db 812 ELPGSOH 818

Mon Jul 8 07:51:46 2002

us-09-786-648-2.ispt

Page 6

Search completed: July 3, 2002, 09:06:02
Job time: 616 sec